

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OMprotein - protein search, using sw model

Ran on: May 24, 2004, 19:05:42 ; Search time 51 Seconds
(without alignments)
626.037 Million cell updates/sec

Title: US-09-756-541-1
Perfect score: 561
Sequence: 1 FYKVEAIVRPWRIQQVSSAL.....VLPVSDVIRVTGERGEKAE 113

Scoring table: 3LQSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 3
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	561	100.0	113	4	AAB69495 Arabidops
2	509	90.7	113	4	AAB69496 Ricinus c
3	312.5	55.7	112	4	AAB69503 Synchoco
4	308.5	55.0	121	6	ABP79052 N. gonorr
5	287.5	51.2	112	4	AAB69501 Azospiril
6	284.5	50.7	112	4	AAB69497 Klebsiell
7	284.5	50.7	112	4	AAB69498 Escherich
8	278.5	49.6	112	4	AAG93226 C glutami
9	264.5	47.1	112	4	AAB69502 Rhodobact
10	260.5	45.4	124	6	ADA34150 Acinetoba
11	257	45.8	111	4	AAB69499 Rhizobium
12	256.5	45.7	160	4	ABG24254 Novel hum
13	256	45.6	111	4	AAB69500 Bradyrhiz
14	246.5	43.9	89	4	AAB79181 Corynebact
15	225	40.1	112	5	ABB54937 Lactococc
16	217.5	38.8	91	5	ABP03233 Human ORF
17	205.5	36.6	121	5	ABB49279 Listeria
18	203	36.2	243	4	ABG24253 Novel hum
19	203	36.2	579	4	ABG24250 Novel hum
20	183	32.6	111	4	AAB69505 Methanoco
21	162.5	29.0	1285	4	ABG26219 Novel hum
22	155.5	27.7	105	4	AAB69504 Methanoco
23	125	22.3	736	4	ABG24252 Novel hum
24	80.5	14.3	1036	6	ABU49254 Protein e
25	74.5	13.3	154	2	AAR05799 PHOMP-2 g

26	73	13.0	639	4	ABG21490	Novel hum
27	72.5	12.9	141	6	ABU25143	Protein e
28	72.5	12.9	1227	2	AAAY31990	Chimeric
29	72	12.8	109	6	ADB12369	Alloiococ
30	72	12.8	340	2	AAR14528	C-termina
31	72	12.8	340	2	AAR41019	Insectici
32	72	12.8	340	2	AAR75409	B. sphaer
33	71.5	12.7	1208	4	AAU02093	Bacillus
34	71.5	12.7	1209	4	AAU02094	Bacillus
35	71.5	12.7	1217	4	AAU02092	Bacillus
36	71.5	12.7	1228	2	AAR50955	Bacillus
37	71	12.7	267	5	ABP66214	Bifidobac
38	70.5	12.6	282	2	AAW73375	S. aureus
39	70.5	12.6	282	6	ABM71198	Staphyloc
40	70.5	12.6	425	6	ABU06030	N. mening
41	70.5	12.6	1292	6	ABU35653	Protein e
42	70	12.5	377	6	ABM65566	Propionib
43	70	12.5	401	4	AAU44161	Propionib
44	70	12.5	401	6	ABM40680	Propionib
45	70	12.5	19938	6	ABP76681	Streptomy

ALIGNMENTS

RESULT 1
AAB69495
ID AAB69495 standard; protein; 113 AA.
XX
AC AAB69495;
XX
DT 23-APR-2001 (first entry)
XX
DE Arabidopsis thaliana PII protein.
XX
KW Arabidopsis thaliana; PII; plant nitrogen regulatory gene; P-PII;
nitrogen assimilation; transgenic plant; herbicide screening.
XX
OS Arabidopsis thaliana.
XX
PN US6177275-B1.
XX
PD 23-JAN-2001.
XX
PF 23-JUL-1997; 97US-00899330.
XX
PR 24-JUL-1996; 96US-0022328P.
XX
(UJNY) UNIV NEW YORK STATE.
Coruzzi GM, Lam H, Hsieh M;
WPI; 2001-158572/16.
N-PSDB; AAF58581, AAF58583.
Novel P-II genes capable of regulating plant nitrogen assimilation,
useful for transgenic plant production, and as probes for isolating
additional genomic clones having P-II gene promoters.
Claim 1; Fig 1; 35pp; English.
The present sequence is encoded by a nitrogen regulatory PII gene. Novel
plant PII (also called P-II) nucleotide sequences have been isolated.
They are useful for regulating nitrogen assimilation in plants, and in
transgenic plant production. They are also used to engineer organisms
that overexpress wild-type or mutant P-II regulatory proteins. P-II
nucleotides are useful for in vitro screening of herbicides. P-II
nucleotides may be used as probes for isolating additional genomic clones
with the promoters of P-II genes. P-II promoters are light- and/or
sucrose-inducible, and are suitable for genetic engineering of plants

Sequence 113 AA;

Query Match 100.0%; Score 561; DB 4; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.8e-62;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYKVEAIVRPWRIQQVSSALLKIGIRGVTWSDVRGFGAQQGSTERHGGSEFSEDKFVAKV 60
DB 1 FYKVEAIVRPWRIQQVSSALLKIGIRGVTWSDVRGFGAQQGSTERHGGSEFSEDKFVAKV 60
QY 61 KMEIVVKDQVESVINTIIEGARTGEIGDGKIFVLPVSDVIRVIRTGERGEKAE 113
DB 61 KMEIVVKDQVESVINTIIEGARTGEIGDGKIFVLPVSDVIRVIRTGERGEKAE 113

RESULT 2
AAB69496
ID AAB69496 standard; protein; 113 AA.

XX AC AAB69496;
XX DT 23-APR-2001 (first entry)
XX DE Ricinus communis PII protein.
XX KW Castor bean; PII; plant nitrogen regulatory gene; P-II;
XX KW nitrogen assimilation; transgenic plant; herbicide screening.
XX OS Ricinus communis.
XX PN US6177275-B1.
XX PD 23-JAN-2001.
XX PF 23-JUL-1997; 97US-00899330.
XX PR 24-JUL-1996; 96US-0022328P.
XX PA (UUNY) UNIV NEW YORK STATE.
XX PI Coruzzi GM, Lam H, Hsieh M;
XX PS WPI; 2001-158572/16.
XX CC 23-JAN-2001.
XX PD 23-JUL-1997; 97US-00899330.
XX PR 24-JUL-1996; 96US-0022328P.
XX PA (UUNY) UNIV NEW YORK STATE.
XX PI Coruzzi GM, Lam H, Hsieh M;
XX PS WPI; 2001-158572/16.
XX DR N-PSDB; AAF58582, AAF58584.

Novel P-II genes capable of regulating plant nitrogen assimilation,
useful for transgenic plant production, and as probes for isolating
additional genomic clones having P-II gene promoters.

Claim 1; Fig 1; 35pp; English.

The present sequence is encoded by a nitrogen regulatory PII gene. Novel
plant PII (also called P-II) nucleotide sequences have been isolated.
They are useful for regulating nitrogen assimilation in plants, and in
transgenic plant production. They are also used to engineer organisms
that overexpress wild-type or mutant P-II regulatory proteins. P-II
proteins are useful for in vitro screening of herbicides. P-II
nucleotides may be used as probes for isolating additional genomic clones
with the promoters of P-II genes. P-II promoters are light- and/or
sucrose-inducible, and are suitable for genetic engineering of plants

Sequence 113 AA;

Query Match 90.7%; Score 509; DB 4; Length 113;
Best Local Similarity 90.3%; Pred. No. 6.3e-56;
Matches 102; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 FYKVEAIVRPWRIQQVSSALLKIGIRGVTWSDVRGFGAQQGSTERHGGSEFSEDKFVAKV 60
DB 1 FYKVEAIVRPWRIQQVSSALLKIGIRGVTWSDVRGFGAQQGSTERHGGSEFSEDKFVAKV 60
QY 61 KMEIVVKDQVESVINTIIEGARTGEIGDGKIFVLPVSDVIRVIRTGERGEKAE 113
DB 61 KMEIVVKDQVEDVIEKIIIEARTGEIGDGKIFLLPVSDVIRVIRTGERGDKAE 113

RESULT 3
AAB69503
ID AAB69503 standard; protein; 112 AA.
XX AC AAB69503;
XX DT 23-APR-2001 (first entry)
XX DE Synecchococcus strain PCC 7942 PII protein.
XX KW Synecchococcus; PII; plant nitrogen regulatory gene; P-II;
XX KW nitrogen assimilation; transgenic plant; herbicide screening.

Synechococcus sp.

US6177275-B1.

23-JAN-2001.

23-JUL-1997; 97US-00899330.

24-JUL-1996; 96US-0022328P.

(UUNY) UNIV NEW YORK STATE.

Coruzzi GM, Lam H, Hsieh M;

WPI; 2001-158572/16.

Novel P-II genes capable of regulating plant nitrogen assimilation,
useful for transgenic plant production, and as probes for isolating
additional genomic clones having P-II gene promoters.

Example; Fig 1; 35pp; English.

The present sequence is encoded by a nitrogen regulatory PII gene. Novel
plant PII (also called P-II) nucleotide sequences have been isolated.
They are useful for regulating nitrogen assimilation in plants, and in
transgenic plant production. They are also used to engineer organisms
that overexpress wild-type or mutant P-II regulatory proteins. P-II
proteins are useful for in vitro screening of herbicides. P-II
nucleotides may be used as probes for isolating additional genomic clones
with the promoters of P-II genes. P-II promoters are light- and/or
sucrose-inducible, and are suitable for genetic engineering of plants

Sequence 112 AA;

Query Match 55.7%; Score 312.5; DB 4; Length 112;
Best Local Similarity 56.4%; Pred. No. 3.7e-31;
Matches 62; Conservative 23; Mismatches 24; Indels 1; Gaps 1;

QY 3 KVEAIVRPWRIQQVSSALLKIGIRGVTWSDVRGFGAQQGSTERHGGSEFSEDKFVAKVM 62

DB 3 KIEAIIREFKLDVKIALVNAGIVGMTVSEVRGFGKQGTERTVSGSEYTVF-FLQKCLKL 61

QY 63 EIVVKDQVESVINTIIEGARTGEIGDGKIFVLPVSDVIRVIRTGERGEKA 112

DB 62 EIVVEDAQDVTVIDKIVAAARTGEIGDGKIFVSPVDQTIIRITGEKNADA 111

RESULT 4
ABP79052
ID ABP79052 standard; protein; 121 AA.

XX AC ABP79052;

XX DT 07-MAR-2003 (first entry)

XX DE N. gonorrhoeae amino acid sequence SEQ ID 4634.

XX KW Antibacterial; infection; vaccine; gene therapy.

OS Neisseria gonorrhoeae.
XX
XX WO200279243-A2.
XX
XX PD 10-OCT-2002.
XX
XX PF 12-FEB-2002; 2002WO-IB002069.
XX
XX PR 12-FEB-2001; 2001GB-00003424.
XX
XX PA (CHIR-) CHIRON SPA.
XX
XX PI Fontana MR, Pizza M, Massignani V, Monaci E;
XX
XX DR WPI; 2003-058415/05.
XX
XX DR N-PSDB; ABZ40022.
XX
XX PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX PT medicament for treating or preventing N. gonorrhoeae infection.
XX
XX PS Disclosure; Page 528; 815pp; English.
XX
XX CC The present invention relates to proteins from Neisseria gonorrhoeae.
XX CC Also disclosed are the nucleic acid molecules encoding the proteins and
XX CC antibodies that specifically bind to the proteins. The composition
XX CC comprising the protein, nucleic acid or antibody is useful for the
XX CC manufacture of a medicament for treating or preventing N. gonorrhoeae
XX CC infection, this may be in the form of a vaccine or gene therapy.
XX CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
XX CC molecules of the invention
XX
XX SQ Sequence 121 AA;

Query Match 55.0%; Score 308.5; DB 6; Length 121;
Best Local Similarity 55.5%; Pred. No. 1.3e-30;
Matches 61; Conservative 24; Mismatches 24; Indels 1; Gaps 1;

QY 3 KVEAIVRPWRIQQVSSALLKIGIRGTVTVSDVRGFGAQQGSTERHGGSEFSEDKFVAKVKM 62
Db 12 KIEAIVKPFKLDVREALTEIGITGMTVSEVKGFGRGKQKGTHTIYRGARYAVD-FLPKVKI 70

QY 63 EIVVKDQVESVINTIIEGARTGEIGDGKIFVLPVSDVIRVTGERGEKA 112
Db 71 ELVLADDAVERAIDVIVEARSGKIGDGKIFVLPVVEAIRITGERSDAA 120

RESULT 5
AAB69501
ID AAB69501 standard; protein; 112 AA.
XX
AC AAB69501;
XX
XX DT 11-SEP-2003 (revised)
XX DT 23-APR-2001 (first entry)
XX
XX DE Azospirillum brasilense PII protein.
XX
XX KW Azospirillum brasilense; PII; plant nitrogen regulatory gene; P-II;
XX KW nitrogen assimilation; transgenic plant; herbicide screening.
XX
XX OS Azospirillum brasilense.
XX
XX PN US6177275-B1.
XX
XX PD 23-JAN-2001.
XX
XX PF 23-JUL-1997; 97US-00899330.
XX
XX PR 24-JUL-1996; 96US-0022328P.
XX
XX PA (UYN) UNIV NEW YORK STATE.
XX
XX PI Coruzzi GM, Lam H, Hsieh M;

XX WPI; 2001-158572/16.
XX
XX PT Novel P-II genes capable of regulating plant nitrogen assimilation,
XX PT useful for transgenic plant production, and as probes for isolating
XX PT additional genomic clones having P-II gene promoters.
XX
XX PS Example; Fig 1; 35pp; English.
XX
XX CC The present sequence is encoded by a nitrogen regulatory PII gene. Novel
XX CC plant PII (also called P-II) nucleotide sequences have been isolated.
XX CC They are useful for regulating nitrogen assimilation in plants, and in
XX CC transgenic plant production. They are also used to engineer organisms
XX CC that overexpress wild-type or mutant P-II regulatory proteins. P-II
XX CC proteins are useful for in vitro screening of herbicides. P-II
XX CC nucleotides may be used as probes for isolating additional genomic clones
XX CC with the promoters of P-II genes. P-II promoters are light- and/or
XX CC sucrose-inducible, and are suitable for genetic engineering of plants.
XX CC (Updated on 11-SEP-2003 to standardise OS field)
XX
XX SQ Sequence 112 AA;

Query Match 51.2%; Score 287.5; DB 4; Length 112;
Best Local Similarity 49.1%; Pred. No. 5.3e-28;
Matches 54; Conservative 27; Mismatches 28; Indels 1; Gaps 1;

QY 3 KVEAIVRPWRIQQVSSALLKIGIRGTVTVSDVRGFGAQQGSTERHGGSEFSEDKFVAKVKM 62
Db 3 KIEAIVKPFKLDVREALTEIGITGMTVSEVKGFGRGKQKGTHTIYRGARYAVD-FLPKVKI 61

QY 63 EIVVKDQVESVINTIIEGARTGEIGDGKIFVLPVSDVIRVTGERGEKA 112
Db 62 EVVMDSLVERAIEAIIQAAHTGRIGDGKIFVTPVEEVVIRITGEKGGDA 111

RESULT 6
AAB69497
ID AAB69497 standard; protein; 112 AA.
XX
AC AAB69497;
XX
XX DT 23-APR-2001 (first entry)
XX
XX DE Klebsiella pneumoniae PII protein.
XX
XX KW Klebsiella pneumoniae; PII; plant nitrogen regulatory gene; P-II;
XX KW nitrogen assimilation; transgenic plant; herbicide screening.
XX
XX OS Klebsiella pneumoniae.
XX
XX PN US6177275-B1.
XX
XX PD 23-JAN-2001.
XX
XX PF 23-JUL-1997; 97US-00899330.
XX
XX PR 24-JUL-1996; 96US-0022328P.
XX
XX PA (UYN) UNIV NEW YORK STATE.
XX
XX PI Coruzzi GM, Lam H, Hsieh M;
XX
XX DR WPI; 2001-158572/16.
XX
XX PT Novel P-II genes capable of regulating plant nitrogen assimilation,
XX PT useful for transgenic plant production, and as probes for isolating
XX PT additional genomic clones having P-II gene promoters.
XX
XX PS Example; Fig 1; 35pp; English.
XX
XX CC The present sequence is encoded by a nitrogen regulatory PII gene. Novel
XX CC plant PII (also called P-II) nucleotide sequences have been isolated.
XX CC They are useful for regulating nitrogen assimilation in plants, and in

Db 63 VIISDAQAEVINIIVETARTGKVGDKVWMTNIEELVRVRTGERGEAA 111

RESULT 9

AAB69502

ID AAB69502 standard; protein; 112 AA.

XX AC AAB69502;

XX DT 23-APR-2001 (first entry)

XX DE Rhodobacter capsulatus PII protein.

XX KW Rhodobacter capsulatus; PII; plant nitrogen regulatory gene; P-PII;

XX KW nitrogen assimilation; transgenic plant; herbicide screening.

XX OS Rhodobacter capsulatus.

XX PN US6177275-B1.

XX PD 23-JAN-2001.

XX PF 23-JUL-1997; 97US-00899330.

XX PR 24-JUL-1996; 96US-0022328P.

XX PA (UUNY) UNIV NEW YORK STATE.

XX PI Coruzzi GM, Lam H, Hsieh M;

XX DR WPI; 2001-158572/16.

XX PT Novel P-II genes capable of regulating plant nitrogen assimilation,

XX PT useful for transgenic plant production, and as probes for isolating

XX PT additional genomic clones having P-II gene promoters.

XX PS Example; Fig 1; 35pp; English.

XX CC The present sequence is encoded by a nitrogen regulatory PII gene. Novel

XX CC plant PII (also called P-II) nucleotide sequences have been isolated.

XX CC They are useful for regulating nitrogen assimilation in plants, and in

XX CC transgenic plant production. They are also used to engineer organisms

XX CC that overexpress wild-type or mutant P-II regulatory proteins. P-II

XX CC proteins are useful for in vitro screening of herbicides. P-II

XX CC nucleotides may be used as probes for isolating additional genomic clones

XX CC with the promoters of P-II genes. P-II promoters are light- and/or

XX CC sucrose-inducible, and are suitable for genetic engineering of plants

XX SQ Sequence 112 AA;

Query Match 47.1%; Score 264.5; DB 4; Length 112;

Best Local Similarity 48.2%; Pred. No. 4.2e-25;

Matches 53; Conservative 25; Mismatches 31; Indels 1; Gaps 1;

QY 3 KVEAIVRPWRIQQVSSALLKIGIRGVTVDVRGFGAQQGSTERHGGSEFSEDKFVAKVKM 62

Db 3 KVEAIIKPKFLDEKALQEAGIQGLSVIEVKGFGRQKGHTELRGAEYVVD-FLPKVKI 61

QY 63 EIVVKKDQVESVINTIIEGARTGEIGDGKIFVLPVSDVIRVRTGERGEKA 112

Db 62 EMVLPDEMVDIAIEAIVGAARTEKIGDGKIFVSSIEQAIRIRTGEGEDA 111

RESULT 10

ADA34150

ID ADA34150 standard; protein; 124 AA.

XX AC ADA34150;

XX DT 20-NOV-2003 (first entry)

XX DE Acinetobacter baumannii protein #1311.

KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;

KW plant biocontrol agent.

XX OS Acinetobacter baumannii.

XX PN US6562958-B1.

XX PD 13-MAY-2003.

XX PF 04-JUN-1999; 99US-00328352.

XX PR 09-JUN-1998; 98US-0088701P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX PI Breton G, Bush D;

XX DR WPI; 2003-576092/54.

XX DR N-PSDB; ADA30024.

XX PT New Acinetobacter baumani proteins and nucleic acids, useful as reagents

XX PT for diagnosing a bacterial disease, as components of antibacterial

XX PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for

XX PT plants.

XX PS Example; SEQ ID NO 5437; 328pp; English.

XX CC The invention relates to isolated Acinetobacter baumannii nucleic acids.

XX CC The A. baumannii nucleic acids and polypeptides are useful as reagents

XX CC for diagnosing a bacterial disease, as components of antibacterial

XX CC vaccines, as targets for antibacterial drugs, to detect the presence of

XX CC A. baumannii and other Acinetobacter species in a sample, in screening

XX CC compounds for the ability to interfere with the A. baumannii life cycle

XX CC or to inhibit A. baumannii infection, and as biocontrol agents for

XX CC plants. The present sequence represents the amino acid sequence of an A.

XX CC baumannii protein.

XX SQ Sequence 124 AA;

Query Match 46.4%; Score 260.5; DB 6; Length 124;

Best Local Similarity 48.6%; Pred. No. 1.5e-24;

Matches 53; Conservative 23; Mismatches 32; Indels 1; Gaps 1;

QY 4 VEAIVRPWRIQQVSSALLKIGIRGVTVDVRGFGAQQGSTERHGGSEFSEDKFVAKVKME 63

Db 16 VTAIVKPKFLDDVREALSDIGVQGITVTEVKGFGRHKGHTELRGAEYVVD-FLPKVKIE 74

QY 64 IWKVKDQVESVINTIIEGARTGEIGDGKIFVLPVSDVIRVRTGERGEKA 112

Db 75 IAISDEMVDVAVIESITRVASTGKIGDGKIFVTNLEQVIRIRTGTPDA 123

RESULT 11

AAB69499

ID AAB69499 standard; protein; 111 AA.

XX AC AAB69499;

XX DT 23-APR-2001 (first entry)

XX DE Rhizobium leguminosarum PII protein.

XX KW Rhizobium leguminosarum; PII; plant nitrogen regulatory gene; P-II;

XX KW nitrogen assimilation; transgenic plant; herbicide screening.

XX OS Rhizobium leguminosarum.

XX PN US6177275-B1.

XX PD 23-JAN-2001.

XX PF 23-JUL-1997; 97US-00899330.

```
PR 24-JUL-1996; 96US-0022328P.
XX (UUNY ) UNIV NEW YORK STATE.
PA Coruzzi GM, Lam H, Hsieh M;
XX WPI; 2001-158572/15.
XX Novel P-P-II genes capable of regulating plant nitrogen assimilation,
PT useful for transgenic plant production, and as probes for isolating
PT additional genomic clones having P-P-II gene promoters.
XX
PS Example; Fig 1; 35pp; English.
XX
CC The present sequence is encoded by a nitrogen regulatory P-II gene. Novel
CC plant P-II (also called P-P-II) nucleotide sequences have been isolated.
CC They are useful for regulating nitrogen assimilation in plants, and in
CC transgenic plant production. They are also used to engineer organisms
CC that overexpress wild-type or mutant P-P-II regulatory proteins. P-P-II
CC proteins are useful for in vitro screening of herbicides. P-P-II
CC nucleotides may be used as probes for isolating additional genomic clones
CC with the promoters of P-P-II genes. P-P-II promoters are light- and/or
CC sucrose-inducible, and are suitable for genetic engineering of plants
XX
SQ Sequence 111 AA;
Query Match 45.8%; Score 257; DB 4; Length 111;
Best Local Similarity 47.7%; Pred. No. 3.7e-24;
Matches 51; Conservative 27; Mismatches 27; Indels 2; Gaps 2;
QY 3 KVEAIVRPRIQQVSSALLKIGIRGVTSDVRGFGAQQGSTERHGGSEFSEDKFVAKVKM 62
Db 3 KIEAIIKPKLDEVRSF-SGVGLQGITVTEAKGFGKRGKGTHTLYRGAEEYVVD-FLPKVKV 60
QY 63 EIVVKDQVESVINTIIEGARTGEIGDKIFVLPVSDVIRVVTGERG 109
Db 61 EVVLADENAEVIEAIRKAAQTGRIGDKIFVSNVBEVIRITGETG 107
RESULT 12
ABG24254
ID ABG24254 standard; protein; 160 AA.
XX
AC ABG24254;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #24245.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US0008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS88441.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
```

```
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 54613; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities in
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 160 AA;
Query Match 45.7%; Score 256.5; DB 4; Length 160;
Best Local Similarity 45.9%; Pred. No. 6.9e-24;
Matches 50; Conservative 28; Mismatches 30; Indels 1; Gaps 1;
QY 4 VEAIVRPRIQQVSSALLKIGIRGVTSDVRGFGAQQGSTERHGGSEFSEDKFVAKVKME 63
Db 52 VTVIIKPKLEDEVREALSSIGIQGLTVEVKGFGKRGKHAELRYGAEYSVN-FLPKVKID 110
QY 64 IVVKDQVESVINTIIEGARTGEIGDKIFVLPVSDVIRVVTGERGEKA 112
Db 111 VAIADDQLDEVIDIVSKAAYTGKIGDKIFVAELQRVIRIRTGEADEAA 159
RESULT 13
AAB69500
ID AAB69500 standard; protein; 111 AA.
XX
AC AAB69500;
XX
DT 23-APR-2001 (first entry)
XX
DE Bradyrhizobium japonicum PII protein.
XX
KW Bradyrhizobium japonicum; PII; plant nitrogen regulatory gene; P-II;
KW nitrogen assimilation; transgenic plant; herbicide screening.
XX
OS Bradyrhizobium japonicum.
XX
PN US6177275-B1.
XX
PD 23-JAN-2001.
XX
PF 23-JUL-1997; 97US-00899330.
XX
PR 24-JUL-1996; 96US-0022328P.
XX
PA (UUNY ) UNIV NEW YORK STATE.
XX
PI Coruzzi GM, Lam H, Hsieh M;
XX
DR WPI; 2001-158572/16.
XX
PT Novel P-P-II genes capable of regulating plant nitrogen assimilation,
PT useful for transgenic plant production, and as probes for isolating
PT additional genomic clones having P-P-II gene promoters.
```

XX Example; Fig 1; 35pp; English.

PS The present sequence is encoded by a nitrogen regulatory PII gene. Novel

XX plant PII (also called P-II) nucleotide sequences have been isolated.

CC They are useful for regulating nitrogen assimilation in plants, and in

CC transgenic plant production. They are also used to engineer organisms

CC that overexpress wild-type or mutant P-II regulatory proteins. P-II

CC proteins are useful for in vitro screening of herbicides. P-II

CC nucleotides may be used as probes for isolating additional genomic clones

CC with the promoters of P-II genes. P-II promoters are light- and/or

CC sucrose-inducible, and are suitable for genetic engineering of plants

XX Sequence 111 AA;

SQ

Query Match 45.6%; Score 256; DB 4; Length 111;

Best Local Similarity 47.7%; Pred. No. 4.9e-24;

Matches 51; Conservative 26; Mismatches 28; Indels 2; Gaps 2;

Qy 3 KVEAIVRPRIQQVSSALLKIGIRGVTVDVVRGFGAOGGSTERHGGSEFSEDKFVAKVKM 62

Db 3 KLEAIIXPKLDEVRS-LSGVGLQGITVTEAKGFGKRGHTDLYRGAEYVD-FLPKVKI 60

Qy 63 EIVVKDQVESVINTIIEGARTGEIGDKIFVLPVSDVIRVTRTGERG 109

Db 61 EIVIGDDLVERAIDAIRRAAQTRIGDGKIFVSNIEAIRIRTGESG 107

RESULT 14

AAB79181

ID AAB79181 standard; protein; 89 AA.

XX AAB79181;

AC AAB79181;

XX

DT 30-APR-2001 (first entry)

XX

DE Corynebacterium glutamicum HA protein sequence SEQ ID NO:318.

XX

KW Corynebacterium glutamicum; homeostasis; adaptation; HA protein;

KW fine chemical production; organic acid; proteinogenic amino acid;

KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;

KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;

KW carbohydrate; aromatic compound; vitamin; cofactor; polypeptide; enzyme;

KW diagnosis; Corynebacterium diphtheriae; genetic engineering;

KW Brevibacterium; environmental condition.

XX

OS Corynebacterium glutamicum.

XX

PN WO200100842-A2.

XX

PD 04-JAN-2001.

XX

PF 23-JUN-2000; 2003WO-IB000911.

XX

PR 25-JUN-1999; 99US-0141031P.

PR 08-JUL-1999; 99DE-01031636.

PR 09-JUL-1999; 99DE-01032125.

PR 09-JUL-1999; 99DE-01032126.

PR 09-JUL-1999; 99DE-01032127.

PR 09-JUL-1999; 99DE-01032128.

PR 09-JUL-1999; 99DE-01032129.

PR 09-JUL-1999; 99DE-01032226.

PR 14-JUL-1999; 99DE-01032920.

PR 14-JUL-1999; 99DE-01032922.

PR 14-JUL-1999; 99DE-01032924.

PR 14-JUL-1999; 99DE-01032928.

PR 14-JUL-1999; 99DE-01032930.

PR 14-JUL-1999; 99DE-01032933.

PR 14-JUL-1999; 99DE-01032935.

PR 14-JUL-1999; 99DE-01032973.

PR 14-JUL-1999; 99DE-01033002.

PR 14-JUL-1999; 99DE-01033003.

PR 14-JUL-1999; 99DE-01033005.

PR 14-JUL-1999; 99DE-01033006.

PR 31-AUG-1999; 99DE-01041378.

PR 31-AUG-1999; 99DE-01041379.

PR 31-AUG-1999; 99DE-01041390.

PR 31-AUG-1999; 99DE-01041391.

PR 03-SEP-1999; 99DE-01042088.

XX (BADI) BASF AG.

PA

XX Pompejus M, Kroeger B, Schroeder H, Zeider O, Haberhauer G;

PI

XX MPI; 2001-061974/07.

DR N-PSDB; AAF71296.

XX

PT New isolated Corynebacterium glutamicum nucleic acid for production or

PT modulation of production of fine chemicals such as amino acids.

PT nucleosides, nucleotides, lipids, fatty acids, carbohydrates, vitamins or

PT enzymes.

XX

PS Claim 20; Page 563; 712pp; English.

XX

CC AAF71138 to AAF71357 encode the Corynebacterium glutamicum homeostasis

CC and adaptation (HA) proteins given in AAB79023 to AAB79242. The C.

CC glutamicum HA genes (I) can be used in vectors for expression in host

CC cells and production of fine chemicals, such as, an organic acid,

CC proteinogenic or nonproteinogenic amino acid (preferred), purine or

CC pyrimidine base, nucleoside, nucleotide, lipid, saturated or unsaturated

CC fatty acid, diol, carbohydrate, aromatic compound, vitamin, cofactor,

CC polypeptide or enzyme. The amino acids produced can be lysine, glutamine,

CC glutamate, alanine, aspartate, glycine, serine, threonine, methionine,

CC cysteine, valine, leucine, isoleucine, arginine, proline, histidine,

CC tyrosine, phenylalanine, or tryptophan. The fine chemical production can

CC be modulated. The presence of (I) or HA proteins encoded by then are used

CC for diagnosing the presence or activity of Corynebacterium diphtheriae.

CC (I) can be used to map the C. glutamicum genome or can be used as markers

CC for genetically engineered Corynebacterium or Brevibacterium. The HA

CC proteins encoded by the (I) are used to maintain homeostasis in C.

CC glutamicum or help the microorganism to adapt to different environmental

CC conditions

XX

SQ Sequence 89 AA;

Query Match 43.9%; Score 246.5; DB 4; Length 89;

Best Local Similarity 50.6%; Pred. No. 5.7e-23;

Matches 45; Conservative 24; Mismatches 19; Indels 1; Gaps 1;

Qy 24 GIRGVTVDVVRGFGAOGGSTERHGGSEFSEDKFVAKVKMEIVVKDQVESVINTIIEGAR 83

Db 1 GVQGMTVTETQSGQQKGHTEVYRGAEYAVD-FVPKVKIEVIISDAQAEVINIIVETAR 59

Qy 84 TGEIGDGKIFVLPVSDVIRVTRTGERGEKA 112

Db 60 TGVGDGKVVMTNIBELVRVTRTGERGEAA 89

RESULT 15

ABB54937

ID ABB54937 standard; protein; 112 AA.

XX

AC ABB54937;

XX

DT 29-AUG-2003 (revised)

DT 16-MAY-2002 (first entry)

XX

DE Lactococcus lactis protein glbB.

XX

KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.

XX

OS Lactococcus lactis; IL1403.

XX

PN FR2807446-A1.

XX

PD 12-OCT-2001.

```

XX 11-APR-2000; 2000FR-00004530.
PF
XX 11-APR-2000; 2000FR-00004530.
PR
XX 11-APR-2000; 2000FR-00004530.
XX
XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA
XX
XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;
PI
XX WPI; 2002-043418/06.
DR
XX
XX New nucleotide sequence useful in the identification or Lactococcus
PT lactis and related species.
PT
XX
XX Claim 6; SEQ ID NO 1639; 2504pp; French.
PS
XX
XX The present invention is related to a Lactococcus lactis nucleotide
CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The nucleic
CC acid sequence is useful in the detection and/or amplification of nucleic
CC acid sequence, particularly to identify Lactococcus lactis or related
CC species. The proteins of the invention are useful for the biosynthesis or
CC biodegradation of a composition of interest. The invention helps research
CC in lactic bacteria, particularly useful in the production of yogurt and
CC cheese. Note: The sequence data for this patent is based on equivalent
CC patent WO200177334 (published 18-OCT-2001) which is available in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 29-AUG-2003 to
CC standardise OS field)
XX
SQ Sequence 112 AA;

Query Match 40.1%; Score 225; DB 5; Length 112;
Best Local Similarity 43.2%; Pred. No. 4e-20;
Matches 48; Conservative 25; Mismatches 36; Indels 2; Gaps 2;

QY 3 KVEAIVRPWRIQQVSSALLKIG-IRGVTVSDVRGFGAQQGSTERHGGSEFSEDKFVAKVK 61
Db 3 KIEAIIRTDKLEDLKALSDNGLVHGMTVSQVLGYGEQKGFTYVRGQRI-ETLLSLKLK 61
QY 62 MEIVVKQQVESVINTTIEGARTGEIGDKIFVLVPSDVIRVRTGERGEKA 112
Db 62 IEIVSIDEKVDDIVNVIIKAVOTGEVGDGKIFIQFVERVIRIRTSEEDAQA 112

```

Search completed: May 24, 2004, 19:13:02
Job time : 53 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 19:10:08 ; Search time 14 Seconds
(without alignments)
776.403 Million cell updates/sec

Title: US-09-756-541-1
Perfect score: 561
Sequence: 1 FYKVEAIVRPWRIQQVSSAL.....VLPVSDVIRVTGERGEKAE 113

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	561	100.0	196	2 D85024	P II nitrogen sens
2	327.5	58.4	112	2 F70310	nitrogen regulator
3	318.5	56.8	112	2 AH2095	nitrogen regulator
4	308.5	55.0	112	2 H81961	nitrogen regulator
5	307.5	54.8	112	2 B81019	nitrogen regulator
6	306.5	54.6	112	2 A39696	nitrogen regulator
7	289.5	51.6	114	2 D82102	nitrogen regulator
8	287.5	51.2	112	2 S13078	nitrogen regulator
9	284.5	50.7	112	1 RGECP2	nitrogen regulator
10	284.5	50.7	112	2 S04377	nitrogen regulator
11	284.5	50.7	112	2 C91056	hypothetical prote
12	284.5	50.7	112	2 AF0354	nitrogen regulator
13	284.5	50.7	112	2 AH0826	nitrogen regulator
14	282.5	50.4	112	2 C64307	probable nitrogen
15	281.5	50.2	112	2 G85900	hypothetical prote
16	279.5	49.8	112	2 S73175	nitrogen regulator
17	276.5	49.3	112	2 S52328	nitrogen regulator
18	276.5	49.3	112	2 S33180	nitrogen regulator
19	276.5	49.3	112	2 G64467	nitrogen regulator
20	275.5	49.1	112	2 AB2794	nitrogen regulator
21	275.5	49.1	157	2 A97573	nitrogen regulator
22	274.5	48.9	112	2 F64062	nitrogen regulator
23	274.5	48.9	112	2 AD3374	nitrogen regulator
24	273.5	48.8	112	2 C87493	nitrogen regulator
25	272.5	48.6	112	2 G70747	probable nitrogen
26	272.5	48.6	121	2 B82302	nitrogen regulator
27	262.5	46.8	112	2 T35668	nitrogen regulator
28	261.5	46.6	115	2 D69188	nitrogen regulator
29	260.5	46.4	112	2 B69468	nitrogen regulator

30	259.5	46.3	112	2 AB0560	nitrogen regulator
31	257.5	45.9	112	2 D82985	nitrogen regulator
32	257.5	45.9	112	2 AG0381	nitrogen regulator
33	257	45.8	111	2 B26567	nitrogen regulator
34	256.5	45.7	112	2 B64775	nitrogen regulator
35	256.5	45.7	112	2 H90691	nitrogen regulator
36	256.5	45.7	112	2 D85542	nitrogen regulator
37	256	45.6	111	2 A33600	nitrogen regulator
38	255.5	45.5	112	2 H82631	nitrogen regulator
39	251.5	44.8	115	2 B69188	nitrogen regulator
40	250.5	44.7	116	2 AD2915	nitrogen regulator
41	250.5	44.7	116	2 G97689	glnK protein (AJ00
42	245.5	43.8	85	2 S76404	nitrogen regulator
43	244.5	43.6	112	2 D87415	nitrogen regulator
44	238.5	42.5	114	2 H87313	nitrogen regulator
45	225	40.1	113	2 G86823	nitrogen regulator

ALIGNMENTS

RESULT 1

D85024

P II nitrogen sensing protein GLB I [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001

C;Accession: D85024

R;anonymus, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin Nature 402, 769-777, 1999

A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A;Reference number: A85001; MUID:20083488; PMID:10617198

A;Accession: D85024

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-196 <STO>

A;Cross-references: GB:NC_001268; NID:g7268574; PIDN:CAB80683.1; GSPDB:GN00140

C;Genetics:

A;Gene: AT4g01900

A;Map position: 4

Query Match 100.0%; Score 561; DB 2; Length 196;
Best Local Similarity 100.0%; Pred. No. 6.8e-48;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYKVEAIVRPWRIQQVSSALLKIGIRGVTVDVSGFGAQQGGSTERHGGSEFSEDKFVAKV 60

Db 74 FYKVEAIVRPWRIQQVSSALLKIGIRGVTVDVSGFGAQQGGSTERHGGSEFSEDKFVAKV 133

QY 61 KMEIVVKDQVESVINTIIEGARTGEIGDGKIFVLPVSDVIRVTGERGEKAE 113

Db 134 KMEIVVKDQVESVINTIIEGARTGEIGDGKIFVLPVSDVIRVTGERGEKAE 186

RESULT 2

F70310

nitrogen regulatory protein P-II - Aquifex aeolicus

C;Species: Aquifex aeolicus

C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 26-Aug-1999

C;Accession: F70310

R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov V.

Nature 392, 353-358, 1998

A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A;Reference number: A70300; MUID:98196666; PMID:9537320

A;Accession: F70310

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-112 <AQF>

A;Cross-references: GB:AE000674; NID:g2982850; PIDN:AAC06473.1; PID:g2982852; GB:AE000065

A;Experimental source: strain VF5

C;Genetics:

A;Gene: glnB

C;Superfamily: regulatory protein P-II

A;Reference number: A56817; MUID:91355213; PMID:1653017
A;Accession: F56817
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-40 <LI1>
A;Experimental source: PCC 6301
C;Genetics:
A;Gene: glnB
C;Function:
A;Description: involved in signal transduction
A;Note: modification state depends on the nitrogen source and the spectral light quality
C;Superfamily: regulatory protein P-II
C;Keywords: phosphoprotein; signal transduction
F;49/Binding site: phosphate (Ser) (covalent) #status predicted
F;51/Binding site: UMP (Tyr) (covalent) #status predicted

Query Match 54.6%; Score 306.5; DB 2; Length 112;
Best Local Similarity 55.5%; Pred. No. 4e-23;
Matches 61; Conservative 23; Mismatches 25; Indels 1; Gaps 1;

QY 3 KVEAIVRPWRIQQVSSALLKIGIRGVTVDVRGFGAQQGSTERHGGSEFSEDKFVAKVKM 62

Db 3 KIEAIIKPFKLDDEVKIALVNAGIVGTVSEVRGFGKQKQTERYRSGSEYTVF-FLQKLKL 61

QY 63 EIVVKDQVESVINTIIEGARTGEIGDKIFVLPVSDVIRVRTGERGEKA 112

Db 62 EIVVEDAQDVTIDKIVAAARPEIGDKIFVSPVDQTIRIRTKGNADA 111

RESULT 7

D82102

nitrogen regulatory protein P-II [similarity] - Vibrio cholerae (strain N16961 serogroup
C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 24-Aug-2001

C;Accession: D82102

R;Heidelbergl, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: D82102

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-114 <HE1>

A;Cross-references: GB:AE004295; GB:AE003852; NID:99656789; PIDN:AAF95383.1; GSPDB:GN001

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC2239

A;Map position: 1

C;Superfamily: regulatory protein P-II

C;Keywords: phosphoprotein

F;53/Binding site: UMP (Tyr) (covalent) #status predicted

Query Match 51.6%; Score 289.5; DB 2; Length 114;
Best Local Similarity 51.8%; Pred. No. 1.9e-21;
Matches 57; Conservative 24; Mismatches 28; Indels 1; Gaps 1;

QY 3 KVEAIVRPWRIQQVSSALLKIGIRGVTVDVRGFGAQQGSTERHGGSEFSEDKFVAKVKM 62

Db 5 KIEAIIKPFKLDDEVKIALVNAGIVGTVSEVRGFGKQKQTERYRSGSEYTVF-FLQKLKL 63

QY 63 EIVVKDQVESVINTIIEGARTGEIGDKIFVLPVSDVIRVRTGERGEKA 112

Db 64 EIVVTDVADRCVDTIETAGTKIGDKIFITNVERVVRIRTKGEDEDA 113

RESULT 8

SL3078

nitrogen regulatory protein P-II - Azospirillum brasilense

C;Species: Azospirillum brasilense

C;Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999

C;Accession: SL3078

R;de Zamaroczy, M.; Delorme, F.; Elmerich, C.
Mol. Gen. Genet. 224, 421-430, 1990

A;Title: Characterization of three different nitrogen-regulated promoter regions for the
A;Reference number: SL3078; MUID:91094780; PMID:1702507

A;Accession: SL3078

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-112 <ZAM>

A;Cross-references: EMBL:X51499; NID:g38664; PIDN:CAA35867.1; PID:g38665

A;Experimental source: strain SP7

C;Genetics:

A;Gene: glnB

C;Function:

A;Description: P-II protein synthesis is increased under conditions of nitrogen fixation

C;Superfamily: regulatory protein P-II

C;Keywords: phosphoprotein; signal transduction

F;51/Binding site: UMP (Tyr) (covalent) #status predicted

Query Match 51.2%; Score 287.5; DB 2; Length 112;
Best Local Similarity 49.1%; Pred. No. 3e-21;
Matches 54; Conservative 27; Mismatches 28; Indels 1; Gaps 1;

QY 3 KVEAIVRPWRIQQVSSALLKIGIRGVTVDVRGFGAQQGSTERHGGSEFSEDKFVAKVKM 62

Db 3 KIEAIIKPFKLDDEVKIEALHEVGIGKITVTEARGFGRQKQHTELYRGAEYVVD-FLPKVKI 61

QY 63 EIVVKDQVESVINTIIEGARTGEIGDKIFVLPVSDVIRVRTGERGEKA 112

Db 62 EIVMEDSLVERAIEAIQAAHTGRIGDKIFVTPVEEVRIRTKGKGDA 111

RESULT 9

RGECP2

nitrogen regulatory protein P-II.1 - Escherichia coli (strain K-12)

C;Species: Escherichia coli

C;Date: 30-Jun-1988 #sequence_revision 10-Nov-1995 #text_change 01-Mar-2002

C;Accession: C49940; SL5991; A29307; S37753; H65032; S31961

R;Liu, J.; Magasanik, B.

J. Bacteriol. 175, 7441-7449, 1993

A;Title: The glnB region of the Escherichia coli chromosome.

A;Reference number: A49940; MUID:94042920; PMID:8226691

A;Accession: C49940

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-112 <LIU>

A;Cross-references: GB:S67014; NID:g4555660; PIDN:AAB28779.1; PID:g4555663

A;Note: sequence extracted from NCBI backbone (NCBI:139878, NCBI:139882)

R;Vasudevan, S.G.; Armarego, W.L.F.; Shaw, D.C.; Lilley, P.E.; Dixon, N.E.; Poole, R.K.

Mol. Gen. Genet. 226, 49-58, 1991

A;Title: Isolation and nucleotide sequence of the hmp gene that encodes a haemoglobin-li

A;Reference number: SL5991; MUID:91238719; PMID:2034230

A;Accession: SL5991

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-112 <VAS>

A;Cross-references: GB:X58872; NID:g41730; PIDN:CAA41683.1; PID:g41732

A;Experimental source: strain K-12

R;Son, H.S.; Rhee, S.G.

J. Biol. Chem. 262, 8690-8695, 1987

A;Title: Cascade control of Escherichia coli glutamine synthetase: purification and prop

A;Reference number: A29307; MUID:87250488; PMID:2885322

A;Accession: A29307

A;Molecule type: DNA

A;Residues: 1-18,'R',20-80,'E',83-102,'VP' <SON>

A;Cross-references: GB:M16778; NID:g146165; PIDN:AAA23883.1; PID:g146166

R;van Heeswijk, W.C.; Rabenberg, M.; Westerhoff, H.V.; Kahn, D.

Mol. Microbiol. 9, 443-457, 1993

A;Title: The genes of the glutamine synthetase adenylation cascade are not regulated b

A;Reference number: S36254; MUID:94018640; PMID:8412694

A;Accession: S37753

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-12 <VAN>

DB 3 KIDAIIPFKLDDVREALAEVGITGMTVTETVKGFGRQKQHTELYRGAAYMVD-FLEPKVKI 61

QY 63 EIVVKDQVESVINTTIEGARTGEIGDGKIFVLPSVDVIRVRTGERGEKA 112
||| | : : : ||| ||| ||| : : : : :
62 EIVVADDIVDTCVEALMOTAHTKSIGDGKIFVFDVSRRVRIRTGEQDEEA 111
||| | : : : ||| ||| ||| : : : : :

DB

RESULT 13
AH0826

nitrogen regulatory protein p-II STY2808 [imported] - Salmonella enterica subsp. enteric
C/Species: Salmonella enterica subsp. enterica serovar Typhi
C/Note: this species has also been called Salmonella typhi
C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C/Accession: AH0826
R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A/Reference number: AB0502; MUID:21534947; PMID:11677608
A/Accession: AH0826

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-112 <PAR>

A;Cross-references: GB:AL513382; PIDN:CAD02764.1; PID:q16503774; GSPDB:GN00176

C;Genetics:
A;Gene: STY2808

C;Superfamily: regulatory protein P-II

Query Match	50.7%;	Score 284.5;	DB 2;	Length 112;
Best Local Similarity	50.9%;	Pred. No. 5.9e-21;		
Matches 56;	Conservative 27;	Mismatches 26;	Indels 1;	Gaps 1;

Qy	3	KVEAIVRPWRIQQVSSALLKIGIRGVTVDYRGFGAQQGSTERHGGSEFSDKFVAKVQM	62
Db	3	KIDAIIKPFKLDDVREALAEVGITGMTVTEVXGFGFGKHGTLYRGAETMYD-FLEPKVKZ	61

Qy 63 EIVVKDQVESVINTIIEGARTGEIGDKIFVLPUVDWIRVRTGEREKA 112
 ||||| | : : : ||||| : ||||| ||||| | : ||||| : |||||
Dd 62 EIVVPDDIVDTCTVDTIIETAGTGKIGDGKIFVEDVARFVIRIRTGEEDDA 111
 ||||| | : : : ||||| : ||||| ||||| | : ||||| : |||||

RESULT 14
C64307

probable nitrogen regulatory protein P-II - *Methanococcus jannaschii*
C:Species: *Methanococcus jannaschii*
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: C64307
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
reson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: C64307

A; Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-112 <BUL>

A;Cross-references: GB:U67464; GB:L77117; NID:g1590852; PIDN:AAB98041.1; PID:g1592259; T
C;Genetics:

A;Map position: FOR58979-59317

C;Superfamily: regulatory protein p-II

C:Keywords: phosphoprotein; signal transduction

F:51/Binding site: UMP (Tyr) (covalent) #status predicted

Query Match	50.4%	Score 282.5;	DB 2;	Length 112;
Best Local Similarity	53.6%	Pred. NO. 9.2e-21;		
Matches 59;	Conservative 21;	Mismatches 29;	Indels 1	

QY 3 KVEAIVRPNRIOQVSSALLKIGIRGVTSDVRGFGAQQGSTERHGGSEFSEDKFVAKVKM 62

D _b	3	KVEAIIRPEKLEIVKKALSDAGYVGMTVSEVKRGVQGGIVERYRGREYIVD-LIPKVKI	61
Q _y	63	EIVVKKQOVESVINTIIEGARTGEIGDGKIFVLPVSDVIRVRTGERGEKA	112
D _b	62	ELVVKKEEDVNDVIDIICENARTGNPGDGKIFVIPERVYRVRTGKEEGKEA	111

RESULT 15
C85900

C;Species: *Escherichia coli*
 C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C;Accession: G85900
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A;Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: G85900

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-112 <STO>

A;Cross-references: GB:AE005174; NID:g12516963; PIDN:AAG57667.1; GSPDB:GN00145; UWGP:Z38
A;Experimental source: strain O157:H7, substrain EDL933

C; Genetics:

A;Gene: q1nB

C;Superfamily: regulatory protein P-II

Query Match	50.2%;	Score 281.5;	DB 2;	Length 112;
Best Local Similarity	50.9%;	Pred. No. 1.2e-20;		
Matches	56;	Conservative	26;	Mismatches 27;
				Indels 1

Qy	3	KVEAIVRPNRIQQVSSALLKIGIRGVTVDVRGFQAQGGSTERHGGSEFSDKFVAVKVM	62
	::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :::		
Db	3	KIDAIIPKPLDDVREALAEVGITGMTVTVEKGFGRQKGHTELXRGAEVMD-FLPKVKI	61

Qy	63	EIVVKDQVESVINTLIEGARTGEIGDGKIFVLPVSDVIRVTRGEGEKA	112
		: : : : : : : : : : : : : : :	
Db	62	EIVPDDIVDTCVTLIRTAQTGKIGDGKIFVDFVARVIRTRGEEDEAA	111

Search completed: May 24, 2004, 19:15:40
Job time : 15 secs


```

CC      -!- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE GS
CC      GENE (GLNA). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO
CC      NR-I-PHOSPHATE, THE TRANSCRIPTIONAL ACTIVATOR OF GLNA. WHEN P-II
CC      IS PHOSPHORYLATED, THESE EVENTS ARE REVERSED (BY SIMILARITY).
CC      -!- FUNCTION: IN NITROGEN-LIMITING CONDITIONS, WHEN THE RATIO OF GLN
CC      TO 2-KETOGLUTARATE DECREASES, P-II IS PHOSPHORYLATED WHICH ALLOWS
CC      THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS), THUS ACTIVATING
CC      THE ENZYME (BY SIMILARITY).
CC      -!- SUBUNIT: Homotrimer (By similarity).
CC      -!- PTM: PHOSPHORYLATION DEPENDENT ON THE NITROGEN SOURCE AND SPECTRAL
CC      LIGHT QUALITY (BY SIMILARITY).
CC      -!- SIMILARITY: Belongs to the P(II) protein family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; X97496; CA56127.1; --
CC      EMBL; D90915; BAA18533.1; ALT_INIT.
CC      HSSP; P05826; 2PII.
CC      InterPro; IPR002187; PII_glnB.
CC      InterPro; IPR002332; PII_glnB_UMP_S.
CC      Pfam; PF00543; P-II; 1.
CC      PRINTS; PR00340; PIIGLNB.
CC      ProDom; PD001194; PII_glnB; 1.
CC      PROSITE; PS00496; PII_GLNB_UMP; 1.
CC      PROSITE; PS00638; PII_GLNB_CTER; 1.
CC      Transcription regulation; Nitrogen fixation; Phosphorylation;
CC      Complete proteome.
CC      MOD_RES 49 49 PHOSPHORYLATION (PROBABLE).
CC      SEQUENCE 112 AA; 12397 MW; F9ABD0F5C173B799 CRC64;
CC      -----
CC      Query Match 55.7%; Score 312.5; DB 1; Length 112;
CC      Best Local Similarity 54.5%; Pred. No. 8.2e-25;
CC      Matches 60; Conservative 26; Mismatches 23; Indels 1; Gaps 1;
CC      -----
CC      QY 3 KVEAIVRPWRIQQVSSALLKIGIRGVTVDVVGFGAQQGSTERHGGSEFSEDKFVAKVKM 62
CC      Db 3 KVEAIRPFKLDVKIALVNAGIVGMTVSEVRGFGKQKQTERVGRGSEYTVF-FLQKLKI 61
CC      -----
CC      QY 63 EIVVKDQVESVINTIIEGARTGEIGDKIFVLPVSDVIRVIRTGERGEKA 112
CC      Db 62 EIVVDEGQVDMVVDKLVSAARTGEIGDKIFISPVDVSVIRIRTGKDEA 111
CC      -----
CC      RESULT 6
CC      GLNB_PASMU
CC      ID GLNB_PASMJ STANDARD; PRT; 112 AA.
CC      AC Q9CUK1;
CC      DT 16-OCT-2001 (Rel. 40, Created)
CC      DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC      DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC      DE Nitrogen regulatory protein p-II.
CC      GN GLNB OR PM2004.
CC      OS Pasteurella multocida.
CC      OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
CC      OC Pasteurellaceae; Pasteurella.
CC      OX NCBI_TaxID=747;
CC      RN [1]
CC      RP SEQUENCE FROM N.A.
CC      RC STRAIN=Pm70;
CC      RX MEDLINE=21145866; PubMed=11248100;
CC      RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
CC      RT "Complete genomic sequence of Pasteurella multocida Pm70.";
CC      RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC      -!- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE
CC      GLUTAMINE SYNTHETASE GENE (GLNA). P-II PREVENTS NR-II CATALYZED
CC      CONVERSION OF NR-I TO NR-I-PHOSPHATE, THE TRANSCRIPTIONAL
CC      ACTIVATOR OF GLNA. WHEN P-II IS URIDYLYLATED TO P-II-UMP, THESE

```

```

CC      EVENTS ARE REVERSED. WHEN THE RATIO OF GLN TO 2-KETOGLUTARATE
CC      DECREASES, P-II IS URIDYLYLATED TO P-II-UMP, WHICH CAUSES THE
CC      DEADENYLATION OF GLUTAMINE SYNTHETASE BY GLNE, SO ACTIVATING THE
CC      ENZYME (BY SIMILARITY).
CC      -!- SUBUNIT: Homotrimer (By similarity).
CC      -!- PTM: Uridylylated/deuridylylated by glnd (By similarity).
CC      -!- SIMILARITY: Belongs to the P(II) protein family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; AE006236; AAK04088.1; --
CC      HSSP; P05826; 2PII.
CC      InterPro; IPR002187; PII_glnB.
CC      InterPro; IPR002332; PII_glnB_UMP_S.
CC      Pfam; PF00543; P-II; 1.
CC      PRINTS; PR00340; PIIGLNB.
CC      ProDom; PD001194; PII_glnB; 1.
CC      PROSITE; PS00496; PII_GLNB_UMP; 1.
CC      PROSITE; PS00638; PII_GLNB_CTER; 1.
CC      Transcription regulation; Nitrogen fixation; Complete proteome.
CC      BINDING 51 51 UMP (BY SIMILARITY).
CC      SEQUENCE 112 AA; 12684 MW; 0CC9D24651F47642 CRC64;
CC      -----
CC      Query Match 51.8%; Score 290.5; DB 1; Length 112;
CC      Best Local Similarity 53.6%; Pred. No. 1.4e-22;
CC      Matches 59; Conservative 21; Mismatches 29; Indels 1; Gaps 1;
CC      -----
CC      QY 3 KVEAIVRPWRIQQVSSALLKIGIRGVTVDVVGFGAQQGSTERHGGSEFSEDKFVAKVKM 62
CC      Db 3 KIEAIKPFKLDVRESLSDWGITGMTVTEVRGFGKQKQHTELYRGAETMVD-FLPKVKM 61
CC      -----
CC      QY 63 EIVVKDQVESVINTIIEGARTGEIGDKIFVLPVSDVIRVIRTGERGEKA 112
CC      Db 62 EIVVTDEQVDCIEAIMETAQTGKIGDKIFVYDVERVIRIRTGEEENED 111
CC      -----
CC      RESULT 7
CC      GLNB_AZOBR
CC      ID GLNB_AZOBR STANDARD; PRT; 112 AA.
CC      AC P21193;
CC      DT 01-MAY-1991 (Rel. 18, Created)
CC      DT 01-MAY-1991 (Rel. 18, Last sequence update)
CC      DT 01-NOV-1995 (Rel. 32, Last annotation update)
CC      DE Nitrogen regulatory protein P-II.
CC      GN GLNB.
CC      OS Azospirillum brasilense.
CC      OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
CC      OC Rhodospirillaceae; Azospirillum.
CC      OX NCBI_TaxID=192;
CC      RN [1]
CC      RP SEQUENCE FROM N.A.
CC      RC STRAIN=Sp7;
CC      RX MEDLINE=91094780; PubMed=1702507;
CC      RA de Zamaroczy M., Delorme F., Elmerich C.;
CC      RT "Characterization of three different nitrogen-regulated promoter
CC      regions for the expression of glnB and glnA in Azospirillum
CC      brasilense.";
CC      RL Mol. Gen. Genet. 224:421-430(1990).
CC      RN [2]
CC      RP SEQUENCE OF 88-112 FROM N.A.
CC      RC STRAIN=Sp7;
CC      RX MEDLINE=87076765; PubMed=2878685;
CC      RA Bozouklian H., Elmerich C.;
CC      RT "Nucleotide sequence of the Azospirillum brasilense Sp7 glutamine
CC      synthetase structural gene.";
CC      RL Biochimie 68:1181-1187(1986).
CC      -!- FUNCTION: IN NITROGEN-LIMITING CONDITIONS, WHEN THE RATIO OF GLN

```


CC IS URIDYLYLATED TO P-II-UMP, THESE EVENTS ARE REVERSED.
CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- SIMILARITY: Belongs to the P(II) protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X14012; CAA32177.1; -.
DR PIR; S04377; S04377.
DR HSSP; P05826; 2PII.
DR InterPro; IPR002187; PII_glnB.
DR InterPro; IPR002332; PII_glnB_UMP_S.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR00340; PIIGLNB.
DR ProDom; PD001194; PII_glnB; 1.
DR PROSITE; PS00496; PII_GLNB_UMP; 1.
DR PROSITE; PS00638; PII_GLNB_CTER; 1.
KW Transcription regulation; Nitrogen fixation.
FT BINDING 51 51 UMP.
FT MUTAGEN 50 50 E->Y: GLUTAMINE AUXOTROPHY.
SQ SEQUENCE 112 AA; 12429 MW; CCL5D58A2F225507 CRC64;

Query Match 50.7%; Score 284.5; DB 1; Length 112;
Best Local Similarity 50.9%; Pred. No. 5.6e-22;
Matches 56; Conservative 27; Mismatches 26; Indels 1; Gaps 1;

QY 3 KVEAIVRPWRIQQVSSALLKIGIRGVTVSDVRGFGAQQGSTERHGGSEFSEDKFVAKVKM 62
Db 3 KIDAIKPKFLDDVREALAEVGLTGMTVTEVKGFGKQKGTSLYRGAEMVD-FLPKVKI 61

QY 63 EIVVKKDQVESVINTIIEGARTGEIGDKIFVLPVSDVIRVIRTGERGEKA 112
Db 62 EIVWTDIVDTCTVDTIIRTAQTGKIGDKGKIFVFDVARVIRIRTGEEDAA 111

RESULT 10
Y059_METJA
ID_Y059_METJA STANDARD; PRT; 112 AA.
AC Q60381;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical nitrogen regulatory PII-like protein MJ0059.
GN MJ0059.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -!- SIMILARITY: Belongs to the P(II) protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67464; AAB98041.1; -.
DR PIR; C64307; C64307.
DR HSSP; P38504; 1GNK.
DR TIGR; MJ0059; -.
DR InterPro; IPR002187; PII_glnB.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR00340; PIIGLNB.
DR ProDom; PD001194; PII_glnB; 1.
DR PROSITE; PS00638; PII_GLNB_CTER; 1.
KW Hypothetical protein; Transcription regulation; Complete proteome.
FT BINDING 51 51 UMP (BY SIMILARITY).
SQ SEQUENCE 112 AA; 12490 MW; 3DC3F307C32FE9F3 CRC64;

Query Match 50.4%; Score 282.5; DB 1; Length 112;
Best Local Similarity 53.6%; Pred. No. 8.9e-22;
Matches 59; Conservative 21; Mismatches 29; Indels 1; Gaps 1;

QY 3 KVEAIVRPWRIQQVSSALLKIGIRGVTVSDVRGFGAQQGSTERHGGSEFSEDKFVAKVKM 62
Db 3 KVEAIIIRPEKLEIVKKALSADAGYVGMTVSEVKGKRGVGGIVRYRGREYVD-LIPKVKI 61

QY 63 EIVVKKDQVESVINTIIEGARTGEIGDKIFVLPVSDVIRVIRTGERGEKA 112
Db 62 ELYVKEEDVDNVDIICENARTGNPGDKIFVIPVERVVRVIRTKEEGKEA 111

RESULT 11
GLNB_PORPU
ID_GLNBPORPU STANDARD; PRT; 112 AA.
AC P51254;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Nitrogen regulatory protein P-II.
GN GLNB.
OS Porphyra purpurea.
OG Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OX NCBI_TaxID=2787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Avonport;
RA Reith M.E., Munnholland J.;
RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast
RT genome."
RL Plant Mol. Biol. Rep. 13:333-335(1995).
CC -!- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE
CC GLUTAMINE SYNTHETASE GENE (GLNA). P-II PREVENTS NR-II CATALYZED
CC CONVERSION OF NR-I TO NR-I-PHOSPHATE, THE TRANSCRIPTIONAL
CC ACTIVATOR OF GLNA. WHEN P-II IS URIDYLYLATED TO P-II-UMP, THESE
CC EVENTS ARE REVERSED. WHEN THE RATIO OF GLN TO 2-KETOGLUTARATE
CC DECREASES, P-II IS URIDYLYLATED TO P-II-UMP, WHICH CAUSES THE
CC DEADENYLYLATION OF GLUTAMINE SYNTHETASE, SO ACTIVATING THE ENZYME
CC (BY SIMILARITY).
CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- SIMILARITY: Belongs to the P(II) protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U38804; AAC08140.1; -.
DR PIR; S73175; S73175.
DR HSSP; P05826; 2PII.


```

DR InterPro; IPR002187; PII_glnB.
DR InterPro; IPR002332; PII_GlnB_UMP_S.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR00340; PIIGLNB.
DR PRODOM; PD001194; PII_glnB; 1.
DR PROSITE; PS00496; PII_GLNБ_UMP; 1.
DR PROSITE; PS00638; PII_GLNБ_CTER; 1.
KW Transcription regulation; Nitrogen fixation; Chloroplast.
FT BINDING 51 UMP (BY SIMILARITY).
SQ SEQUENCE 112 AA; 12320 MW; 29416AA49FC37A18 CRC64;

Query Match      49.8%; Score 279.5; DB 1; Length 112;
Best Local Similarity 51.8%; Pred. No. 1.8e-21;
Matches 57; Conservative 21; Mismatches 31; Indels 1; Gaps 1;

QY    3 KVEAIVRPWRIQQVSSALLKIGIRGVTVDVRGFAGQGSGSTERHGGSEFSEDKFFAKVM 62
      :|::||::: ||::||::|| |::|| |::||::||::||::||::||::||::|
Dd    3 KIEAIRPFKLNEVKIALVKGIGGMTVVKVSGFRQKGQTERYKGSYSID-IIDKIKI 61

QY    63 EIVVKKDQVESVNTIIEGARTGEIGDGKIFVLVPVSDVIRVRTGERGEKA 112
      ||::||::: ||::||::||::||::||::||::||::||::||::||::||::|
Db    62 EIIVSDCKVNSITEIITKTAKTGEIGDGKIFISDVVEQVIRIRTNDLNSAA 111

```

```

RESULT 12
GLNB_RHORU          STANDARD;      PRT;    112 AA.
AC    Q53044;
DT    01-NOV-1997 (Rel. 35, Created)
DT    01-NOV-1997 (Rel. 35, Last sequence update)
DT    15-DEC-1998 (Rel. 37, Last annotation update)
DE    Nitrogen regulatory protein P-II.
GN    GLNB.
OS    Rhodospirillum rubrum.
OC    Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC    Rhodospirillaceae; Rhodospirillum.
OX    NCBI_TaxID=1085;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=ATCC 11170 / S1;
RX    MEDLINE=96254013; PubMed=8704966;
RA    Johansson M., Nordlund S.;
RT    "Transcription of the glnB and glnA genes in the photosynthetic
RL    bacterium Rhodospirillum rubrum.";
RN    [2]
RP    SEQUENCE FROM N.A.
RC    STRAIN=ATCC 11170 / S1;
RA    Zhang Y., Ludden P.W., Roberts G.P.;
RL    Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC    -!- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE
CC    GLUTAMINE SYNTHETASE GENE (GLN). P-II PREVENTS NR-II CATALYZED
CC    CONVERSION OF NR-I TO NR-I-PHOSPHATE, THE TRANSCRIPTIONAL
CC    ACTIVATOR OF GLN. WHEN P-II IS URIDYLYLATED TO P-II-UMP, THESE
CC    EVENTS ARE REVERSED. WHEN THE RATIO OF GLN TO 2-KETOGLUTARATE
CC    DECREASES, P-II IS URIDYLYLATED TO P-II-UMP, WHICH CAUSES THE
CC    DEADENYLYLATION OF GLUTAMINE SYNTHETASE, SO ACTIVATING THE ENZYME.
CC    -!- SUBUNIT: Homotrimer (By similarity).
CC    -!- SIMILARITY: Belongs to the P(II) protein family.
CC    -----
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    -----
DR    EMBL; X84158; CAAS8963.1; -.
DR    EMBL; AF029703; AAB84167.1; -.
DR    PIR; S52328; S52328.
DR    HSSP; P05826; 2PII.
DR    InterPro; IPR002187; PII glnB.

```

[illegible]

Search completed: May 24, 2004, 19:13:34
Job time : 11.5 secs

Search completed: May 24, 2004, 19:13:34
Job time : 11.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 19:09:38 ; Search time 36 Seconds
(without alignments)
990.377 Million cell updates/sec

Title: US-09-756-541-1
Perfect score: 561
Sequence: 1 FYKVEAIVPWRIQOVSSAL.....VLPVSDVIRVTGERGEKAE 113

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTRMBL25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	561	100.0	196	10 Q9ZST4	Q9zst4 arabidopsis
2	509	90.7	197	10 Q9ZST5	Q9zst5 ricinus com
3	490	87.3	194	10 Q9ARI4	Q9ari4 medicago sa
4	322.5	57.5	112	16 Q7V5R4	Q7v5r4 prochloroco
5	318.5	56.8	112	16 Q9L422	Q9l422 anabaena sp
6	317.5	56.6	112	16 Q8DLA5	Q8dla5 synechococc
7	316.5	56.4	112	2 Q9LAQ3	Q9laq3 synechococc
8	308.5	55.0	112	16 Q9JWC4	Q9jwc4 neisseria m
9	307.5	54.8	112	16 Q9JXK6	Q9jxk6 neisseria m
10	305.5	54.5	112	2 Q8GQS4	Q8gqs4 acetobacter
11	304.5	54.3	112	16 Q7VEK6	Q7vpk6 haemophilus
12	303.5	54.1	112	2 Q9L400	Q9l400 prochloroco
13	303.5	54.1	112	16 Q7V025	Q7v025 prochloroco
14	301.5	53.7	133	16 Q7U0Z1	Q7uuz1 rhodospirell
15	298.5	53.2	112	16 Q7VA51	Q7va51 prochloroco
16	297.5	53.0	112	16 Q7U8Z7	Q7u8z7 synechococc

17	296.5	52.9	112	2 Q9EZQ2	Q9ezq2 azoarcus sp
18	293.5	52.3	114	16 Q87MF6	Q87mf6 vibrio para
19	292.5	52.1	108	2 Q8KUJ2	Q8kuj2 neisseria m
20	292.5	52.1	114	16 Q8DBE2	Q8dbe2 vibrio vuln
21	291.5	52.0	112	16 Q8EBJ6	Q8ebj6 shewanella
22	289.5	51.6	114	16 Q9KFX3	Q9kpx3 vibrio chol
23	286.5	51.1	105	2 Q8KUJ5	Q8kuj5 neisseria m
24	285.5	50.9	112	2 Q31188	Q31188 rhodobacter
25	284.5	50.7	112	16 Q8ZCQ9	Q8zcg9 yersinia pe
26	282.5	50.4	112	2 Q9AMM9	Q9amm9 azospirillu
27	281.5	50.2	112	16 Q8XWX5	Q8xwx5 ralstonia s
28	281.5	50.2	112	16 Q7WJ14	Q7wj14 bordetella
29	281.5	50.2	112	16 Q7WAE1	Q7wae1 bordetella
30	281.5	50.2	112	16 Q7VVG9	Q7vv99 bordetella
31	278.5	49.6	112	16 Q9X705	Q9x705 corynebacte
32	276.5	49.3	112	2 P94125	P94125 azorhizobiu
33	276.5	49.3	112	16 Q8EAS3	Q8eas3 shewanella
34	276.5	49.3	112	17 Q58740	Q58740 methanococc
35	275.5	49.1	112	2 P94852	P94852 herbaspiril
36	275.5	49.1	157	16 Q8UEI8	Q8uei8 agrobacteri
37	274.5	48.9	112	16 Q8YH26	Q8yh26 brucella me
38	274.5	48.9	112	16 Q8G0T2	Q8g0t2 brucella su
39	273.5	48.8	101	2 Q8KHS5	Q8khs5 neisseria m
40	273.5	48.8	112	2 P70731	P70731 azospirillu
41	273.5	48.8	112	16 Q9A6W4	Q9a6w4 caulobacter
42	272.5	48.6	101	2 Q8KUJ0	Q8kuj0 neisseria m
43	272.5	48.6	112	16 Q8FP24	Q8fp24 corynebacte
44	272.5	48.6	121	16 Q9KUB6	Q9kub6 vibrio chol
45	271.5	48.4	100	2 Q8KI77	Q8ki77 neisseria m

ALIGNMENTS

RESULT 1

ID	Q9ZST4	PRELIMINARY;	PRT;	136 AA.
AC	Q9ZST4;			
DT	01-MAY-1999 (TrEMBLrel. 10, Created)			
DT	01-MAY-1999 (TrEMBLrel. 10, Last sequence update)			
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)			
DE	PII protein (P II nitrogen sensing protein GLB I) (A049019001)			
GN	T7B11.16 OR AT4G01900.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi			
OX	NCBI_TaxID=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99030678; PubMed=9811909;			
RA	Hsieh M.H., Lam H.M., van de Loo F.J., Coruzzi G.;			
RT	"A PII-like protein in Arabidopsis: putative role in nitrogen			
RT	sensing.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 95:13965-13970 (1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RA	Huang E.N., Nascimento L., de la Bastide M., Habermann K., Vil M.D.,			
RA	Preston R.R., Spiegel L.A., See L.H., Shah R., Matero A.,			
RA	O'Shaughnessy A., Rodriguez M., Shekher M., Swaby I., Schutz K.,			
RA	Parnell L.D., Dedhia N.N., McCombie W.R.;			
RT	"Arabidopsis thaliana BAC T7B11 from chromosome IV near 10 cM.";			
RL	Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,			
RA	Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,			
RA	Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,			
RA	Mewes H.W., Lemcke K., Mayer K.F.X.;			
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			

RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Shinn P., Bowser L., Carninci P.,
Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF095455; AAC78333.1; -.
DR EMBL; AC007138; AAD22652.1; -.
DR EMBL; AL161493; CAB80683.1; -.
DR EMBL; BT005209; AAO63273.1; -.
DR PIR; D85024; D85024.
DR HSSP; P05826; 2PII.
DR GO; GO:0030234; F:enzyme regulator activity; IEA.
DR GO; GO:0006808; P:regulation of nitrogen utilization; IEA.
DR InterPro; IPR002187; PII_glnB.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR00340; PIIGLNB.
DR ProDom; PD001194; PII_glnB; 1.
DR PROSITE; PS00638; PII_GLNB_CTER; 1.
SQ SEQUENCE 196 AA; 21275 MW; FE740EA66776F157 CRC64;

Query Match 100.0%; Score 561; DB 10; Length 196;
Best Local Similarity 100.0%; Pred. No. 4.5e-48;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYKVEAIVRPWRIQQVSSALLKIGIRGVTVDVRGFGAQQGSTERHGGSEFSEDKFVAKV 60
Db 74 FYKVEAIVRPWRIQQVSSALLKIGIRGVTVDVRGFGAQQGSTERHGGSEFSEDKFVAKV 133

QY 61 KMEIVVKDQVESVINTIIEGARTGEIGDGKIFVLPVSDVIRVTRGERGEKAE 113
Db 134 KMEIVVKDQVESVINTIIEGARTGEIGDGKIFVLPVSDVIRVTRGERGEKAE 186

RESULT 2
Q9ZST5
ID Q9ZST5 PRELIMINARY; PRT; 197 AA.
AC Q9ZST5;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PII protein (Fragment).
CS Ricinus communis (Castor bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids 2; Malpighiales; Euphorbiaceae; Acalyphoideae; Acalyphaceae;
OC Ricinus.
OX NCBI_TaxID=3988;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99030678; PubMed=9811909;
RA Hsieh M.H., Lam H.M., van de Loo F.J., Coruzzi G.;
RT "A PII-like protein in Arabidopsis: putative role in nitrogen
RT sensing.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:13965-13970 (1998).
DR EMBL; AF095454; AAC78332.1; -.
DR HSSP; P05826; 2PII.
DR GO; GO:0030234; F:enzyme regulator activity; IEA.
DR GO; GO:0006808; P:regulation of nitrogen utilization; IEA.
DR InterPro; IPR002187; PII_glnB.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR00340; PIIGLNB.
DR ProDom; PD001194; PII_glnB; 1.
DR PROSITE; PS00638; PII_GLNB_CTER; 1.
FT NON_TER 1
SQ SEQUENCE 197 AA; 21781 MW; 2D18A50CF15406B20 CRC64;

Query Match 90.7%; Score 509; DB 10; Length 197;
Best Local Similarity 90.3%; Pred. No. 7.1e-43;
Matches 102; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 FYKVEAIVRPWRIQQVSSALLKIGIRGVTVDVRGFGAQQGSTERHGGSEFSEDKFVAKV 60
Db 72 FYKVEAILRPWRVSVQVSSALLKIGIRGVTVDVRGFGAQQGSTERHGGSEFSEDKFVAKV 131

QY 61 KMEIVVKDQVESVINTIIEGARTGEIGDGKIFVLPVSDVIRVTRGERGEKAE 113
Db 132 KMEIVVKDQVEDVIEKIIIEARTGEIGDGKIFLLPVSDVIRVTRGERGEKAE 184

RESULT 3
Q9ARI4
ID Q9ARI4 PRELIMINARY; PRT; 194 AA.
AC Q9ARI4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PII protein.
GN GLNB.
OS Medicago sativa (Alfalfa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids 1; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
OX NCBI_TaxID=3879;
RN [1]
RP SEQUENCE FROM N.A.
RA Garcia-Ibáñeta D., Sengupta-Gopalan C.;
RT "Characterization of PII (GLNB) in alfalfa.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY027892; AAK16221.1; -.
DR HSSP; P38504; 1GNK.
DR GO; GO:0030234; F:enzyme regulator activity; IEA.
DR GO; GO:0006808; P:regulation of nitrogen utilization; IEA.
DR InterPro; IPR002187; PII_glnB.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR00340; PIIGLNB.
DR ProDom; PD001194; PII_glnB; 1.
DR PROSITE; PS00638; PII_GLNB_CTER; 1.
SQ SEQUENCE 194 AA; 21413 MW; 23FA623FF4D97450 CRC64;

Query Match 87.3%; Score 490; DB 10; Length 194;
Best Local Similarity 85.0%; Pred. No. 5.5e-41;
Matches 96; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

QY 1 FYKVEAIVRPWRIQQVSSALLKIGIRGVTVDVRGFGAQQGSTERHGGSEFSEDKFVAKV 60
Db 70 FYKVEAILRPWRIPQVSSGLLKMGIRGVTVDVRGFGAQQGSKERQGGSEFSEDNFVAKV 129

QY 61 KMEIVVKDQVESVINTIIEGARTGEIGDGKIFVLPVSDVIRVTRGERGEKAE 113
Db 130 KMEIVVRKDQVEAVINKIMETARTGEIGDGKIFLLPVSDVIRVTRGERGEQAE 182

RESULT 4
Q7V5R4
ID Q7V5R4 PRELIMINARY; PRT; 112 AA.
AC Q7V5R4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nitrogen regulatory protein P-II.
GN GLNB OR PMT1481.
OS Prochlorococcus marinus (strain MIT 9313).
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=74547;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825698; PubMed=12917642;

[illegible]

Query Match 55.0%; Score 308.5; DB 16; Length 112;
Best Local Similarity 55.5%; Pred. No. 3.8e-23;
Matches 61; Conservative 24; Mismatches 24; Indels 1; Gaps 1;
QY 3 KVEAIVRPWRIQQVSSALLKIGIRGVTVDVVRGFGAQQGSTERHGGSEFSEDKFVAKVKM 62
DB 3 KIEAIVKPFKLDVREALTEIGITGMTVSEVKGFGKQKQKHTIYRGAAYVD-FLPKVKI 61
QY 63 EIVVKKDQVESVINTIIEGARTGEIGDKIFVLPVSDVIRVRTGERGEKA 112
DB 62 ELVLADDAVERAIDVIVEVARSGKIGDKIFVLPVVEAIRINTGERADAA 111
RESULT 9
Q9JXK6 PRELIMINARY; PRT; 112 AA.
AC Q9JXK6;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Nitrogen regulatory protein P-II.
GN NMB1995.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=MC58 / Serogroup B;
RC MEDLINE=20175755; PubMed=10710307;
RX Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.; strain
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58.";
RL Science 287:1809-1815(2000).
DR EMBL; AE002548; AAF42322.1; --
DR PIR; B81019; B81019.
DR HSSP; P05826; 2PII.
DR TIGR; NMB1995; --
DR GO; GO:0030234; F:enzyme regulator activity; IEA.
DR GO; GO:0006808; P:regulation of nitrogen utilization; IEA.
DR InterPro; IPR002187; PII_glnB.
DR InterPro; IPR002332; PII_glnB_UMP_S.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR00340; PIIGLNB.
DR ProDom; PD001194; PII_glnB; 1.
DR PROSITE; PS00638; PII_GlnB_CTER; 1.
DR PROSITE; PS00496; PII_GlnB_UMP; 1.
KW Complete proteome.
SQ SEQUENCE 112 AA; 12311 MW; 22963CF2526D4332 CRC64;
Query Match 54.8%; Score 307.5; DB 16; Length 112;
Best Local Similarity 54.5%; Pred. No. 4.8e-23;
Matches 60; Conservative 25; Mismatches 24; Indels 1; Gaps 1;
QY 3 KVEAIVRPWRIQQVSSALLKIGIRGVTVDVVRGFGAQQGSTERHGGSEFSEDKFVAKVKM 62
DB 3 KIEAIVKPFKLDVREALTEIGITGMTVSEVKGFGKQKQKHTIYRGAAYVD-FLPKVKI 61
QY 63 EIVVKKDQVESVINTIIEGARTGEIGDKIFVLPVSDVIRVRTGERGEKA 112
DB 62 ELVLADDAVERAIDVIVEVARSGKIGDKIFVLPVVEAIRINTGERADAA 111
RESULT 10
Q8GQS4 PRELIMINARY; PRT; 112 AA.
ID Q8GQS4

SEQUENCE FROM N.A.
STRAIN=PCC7002;
RA Sakamoto T., Gruber T., Bryant D.A.;
RT "Nucleotide sequence of glnB and bcp genes in Synechococcus sp.
PCC7002.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF120107; AAF63031.1; --
DR HSSP; P05826; 2PII.
DR GO; GO:0030234; F:enzyme regulator activity; IEA.
DR GO; GO:0006808; P:regulation of nitrogen utilization; IEA.
DR InterPro; IPR002187; PII_glnB.
DR InterPro; IPR002332; PII_glnB_UMP_S.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR00340; PIIGLNB.
DR ProDom; PD001194; PII_glnB; 1.
DR PROSITE; PS00638; PII_GlnB_CTER; 1.
DR PROSITE; PS00496; PII_GlnB_UMP; 1.
SQ SEQUENCE 112 AA; 12432 MW; C3F0AA5625E9F870 CRC64;
Query Match 56.4%; Score 316.5; DB 2; Length 112;
Best Local Similarity 54.5%; Pred. No. 6.1e-24;
Matches 60; Conservative 27; Mismatches 22; Indels 1; Gaps 1;
QY 3 KVEAIVRPWRIQQVSSALLKIGIRGVTVDVVRGFGAQQGSTERHGGSEFSEDKFVAKVKM 62
DB 3 KVEAIVKPFKLDVREALTEIGITGMTVSEVKGFGKQKQKHTIYRGAAYVD-FLPKVKI 61
QY 63 EIVVKKDQVESVINTIIEGARTGEIGDKIFVLPVSDVIRVRTGERGEKA 112
DB 62 EIVIDDDQDVAVDKIVAAARTGEIGDKIFISPVQDQIVRIRTKLEA 111
RESULT 8
Q9JWC4 PRELIMINARY; PRT; 112 AA.
AC Q9JWC4;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Nitrogen regulatory protein P-II 1.
GN GLNB OR NMA047.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=Z2491 / Serogroup A / Serotype 4A;
RC MEDLINE=20222556; PubMed=10761919;
RX Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis Z2491.";
RL Nature 404:502-506(2000).
DR EMBL; AL162753; CAB83744.1; --
DR PIR; H81961; H81961.
DR HSSP; P05826; 2PII.
DR GO; GO:0030234; F:enzyme regulator activity; IEA.
DR GO; GO:0006808; P:regulation of nitrogen utilization; IEA.
DR InterPro; IPR002187; PII_glnB.
DR InterPro; IPR002332; PII_glnB_UMP_S.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR00340; PIIGLNB.
DR ProDom; PD001194; PII_glnB; 1.
DR PROSITE; PS00638; PII_GlnB_CTER; 1.
DR PROSITE; PS00496; PII_GlnB_UMP; 1.
KW Complete proteome.
SQ SEQUENCE 112 AA; 12281 MW; CD893CF24A1A485P CRC64;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 19:05:42 ; Search time 51 Seconds
(without alignments)
626.037 Million cell updates/sec

Title: US-09-756-541-2
Perfect score: 557
Sequence: 1 FYKVEAILRPWRVSQVSSAL.....LLPVS DVIRVTGERGDKAE 113

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	557	100.0	113	4	AAB69496 Ricinus c
2	509	91.4	113	4	AAB69495 Arabidops
3	306.5	55.0	112	4	AAB69503 Synechoco
4	303.5	54.5	121	6	ABP79052 N. gonorr
5	281.5	50.5	112	4	AAB69501 Azospiril
6	276.5	49.6	112	4	AAB69497 Klebsiell
7	274.5	49.3	112	4	AAB69498 Escherich
8	270.5	48.6	112	4	AAG93226 C glutami
9	256.5	46.1	124	6	ADA34150 Acinetoba
10	255.5	45.9	112	4	AAB69502 Rhodobact
11	251	45.1	111	4	AAB69499 Rhizobium
12	248	44.5	111	4	AAB69500 Bradyrhiz
13	245.5	44.1	160	4	ABG24254 Novel hum
14	240.5	43.2	89	4	AAB79181 Corynebac
15	218	39.1	112	5	ABB54937 Lactococc
16	215.5	38.7	91	5	ABP01233 Human ORP
17	207.5	37.3	121	5	ABB49279 Listeria
18	183	32.9	243	4	ABG24253 Novel hum
19	183	32.9	579	4	ABG24250 Novel hum
20	178	32.0	111	4	AAB69505 Methanoco
21	159	28.5	1285	4	ABG26219 Novel hum
22	153.5	27.6	105	4	AAB69504 Methanoco
23	120	21.5	736	4	ABG24252 Novel hum
24	84.5	15.2	1036	6	ABU49254 Protein e
25	78	14.0	639	4	ABG21490 Novel hum

26	77	13.8	709	5	ABP30163 Streptoco
27	77	13.8	713	5	ABP25678 Streptoco
28	75.5	13.6	823	6	ABU21334 Protein e
29	74.5	13.4	274	4	ABB69532 Drosophil
30	73.5	13.2	374	2	AAY43982 Horse alc
31	73.5	13.2	374	2	AAY43981 Horse alc
32	73.5	13.2	547	2	AAR67385 Mitochond
33	73.5	13.2	554	5	ABP28230 Streptoco
34	72	12.9	169	5	ABP41194 Human ova
35	72	12.9	187	7	ADE24811 Human kin
36	71.5	12.8	541	6	ABU49927 Protein e
37	71	12.7	773	5	ABB55204 Lactococc
38	70.5	12.7	573	2	AAR04713 Amino aci
39	70.5	12.7	573	2	AAW01657 Human hea
40	70.5	12.7	573	2	AAW14946 Human hea
41	70.5	12.7	573	2	AAW12345 Human hea
42	70.5	12.7	573	2	AAW56120 Protein s
43	70.5	12.7	573	2	AAY23926 Amino aci
44	70.5	12.7	573	3	AAY93334 Amino aci
45	70.5	12.7	573	4	AAE11756 Human hea

ALIGNMENTS

RESULT 1
AAB69496
ID AAB69496 standard; protein; 113 AA.
XX
AC AAB69496;
XX 23-APR-2001 (first entry)
XX Ricinus communis PII protein.
DE Ricinus communis PII protein.
XX
KW Castor bean; PII; plant nitrogen regulatory gene; P-II;
KW nitrogen assimilation; transgenic plant; herbicide screening.
XX
OS Ricinus communis.
XX
PN US6177275-B1.
XX
PD 23-JAN-2001.
XX
PF 23-JUL-1997; 97US-00899330.
XX
PR 24-JUL-1996; 96US-0022328P.
XX (UUNY) UNIV NEW YORK STATE.
XX Coruzzi GM, Lam H, Hsieh M;
DR WPI; 2001-158572/16.
DR N-PSDB; AAF58582, AAF58584.
XX
PT Novel P-II genes capable of regulating plant nitrogen assimilation,
PT useful for transgenic plant production, and as probes for isolating
PT additional genomic clones having P-II gene promoters.
XX
PS Claim 1; Fig 1; 35pp; English.
XX
CC The present sequence is encoded by a nitrogen regulatory PII gene. Novel
CC plant PII (also called P-II) nucleotide sequences have been isolated.
CC They are useful for regulating nitrogen assimilation in plants, and in
CC transgenic plant production. They are also used to engineer organisms
CC that overexpress wild-type or mutant P-II regulatory proteins. P-II
CC proteins are useful for in vitro screening of herbicides. P-II
CC nucleotides may be used as probes for isolating additional genomic clones
CC with the promoters of P-II genes. P-II promoters are light- and/or
CC sucrose-inducible, and are suitable for genetic engineering of plants
XX
SQ Sequence 113 AA;

Db 63 VIISDAQAEVINIIVETARTGKVGDKVMTNIEELVRVRTGERGEAA 111

RESULT 9
ADA34150
ID ADA34150 standard; protein; 124 AA.
XX
AC ADA34150;
XX
DT 20-NOV-2003 (first entry)
XX
DE Acinetobacter baumannii protein #1311.
XX
KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
KW plant biocontrol agent.
XX
OS Acinetobacter baumannii.
XX
PN US6562958-B1.
XX
PD 13-MAY-2003.
XX
PF 04-JUN-1999; 99US-00328352.
XX
PR 09-JUN-1998; 98US-0088701P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton G, Bush D;
XX
DR WPI; 2003-576092/54.
DR N-PSDB; ADA30024.
XX
PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT for diagnosing a bacterial disease, as components of antibacterial
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT plants.
XX
PS Example; SEQ ID NO 5437; 328pp; English.
XX
CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC The A. baumannii nucleic acids and polypeptides are useful as reagents
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of
CC A. baumannii and other Acinetobacter species in a sample, in screening
CC compounds for the ability to interfere with the A. baumannii life cycle
CC or to inhibit A. baumannii infection, and as biocontrol agents for
CC plants. The present sequence represents the amino acid sequence of an A.
CC baumannii protein.
XX
SQ Sequence 124 AA;

Query Match 46.1%; Score 256.5; DB 6; Length 124;
Best Local Similarity 48.6%; Pred. No. 3.8e-22;
Matches 53; Conservative 22; Mismatches 33; Indels 1; Gaps 1;
QY 4 VEAILRPWRVSVSSALLKIGIRGVTVDVRGFGAQQGSTERQGGSEFSEDKFVAKVKME 63
Db 16 VTAIVKPFKLDVREALSDIGVQGITVTEVKGFGGRHKGHTELYRGAETVVD-FLPKVKIE 74
QY 64 IVVSKDQVEDVIEKIIIEARTGEIGDGKIFLLPVSDVIRVIRTGERGDKA 112
Db 75 IASDEMVDVAVIESITRVASTGKIGDGKIFVTNLEQVIRIRTGETGDDA 123

RESULT 10
AAB69502
ID AAB69502 standard; protein; 112 AA.
XX
AC AAB69502;
XX
DT 23-APR-2001 (first entry)
XX

DE Rhodobacter capsulatus PII protein.
XX
KW Rhodobacter capsulatus; PII; plant nitrogen regulatory gene; P-PII;
KW nitrogen assimilation; transgenic plant; herbicide screening.
XX
OS Rhodobacter capsulatus.
XX
PN US6177275-B1.
XX
DT 23-JAN-2001.
XX
PF 23-JUL-1997; 97US-00899330.
XX
PR 24-JUL-1996; 96US-0022328P.
XX
PA (UYNY) UNIV NEW YORK STATE.
XX
PI Coruzzi GM, Lam H, Hsieh M;
XX
DR WPI; 2001-158572/16.
XX
PT Novel P-PII genes capable of regulating plant nitrogen assimilation,
PT useful for transgenic plant production, and as probes for isolating
PT additional genomic clones having P-PII gene promoters.
XX
PS Example; Fig 1; 35pp; English.
XX
CC The present sequence is encoded by a nitrogen regulatory PII gene. Novel
CC plant PII (also called P-PII) nucleotide sequences have been isolated.
CC They are useful for regulating nitrogen assimilation in plants, and in
CC transgenic plant production. They are also used to engineer organisms
CC that overexpress wild-type or mutant P-PII regulatory proteins. P-PII
CC proteins are useful for in vitro screening of herbicides. P-PII
CC nucleotides may be used as probes for isolating additional genomic clones
CC with the promoters of P-PII genes. P-PII promoters are light- and/or
CC sucrose-inducible, and are suitable for genetic engineering of plants
XX
SQ Sequence 112 AA;

Query Match 45.9%; Score 255.5; DB 4; Length 112;
Best Local Similarity 47.3%; Pred. No. 4.4e-22;
Matches 52; Conservative 26; Mismatches 31; Indels 1; Gaps 1;
QY 3 KVEAILRPWRVSVSSALLKIGIRGVTVDVRGFGAQQGSTERQGGSEFSEDKFVAKVKM 62
Db 3 KVEAIIKPFKLDVKEALQEAGIQLSVIEVKGFGGRKHGHTELYRGAETVVD-FLPKVKI 61
QY 63 EIVVSKDQVEDVIEKIIIEARTGEIGDGKIFLLPVSDVIRVIRTGERGDKA 112
Db 62 EMVLPDEMVDIAIEAIVGAARTEKIGDGKIFVSSIEQAIRIRTGETGEDA 111

RESULT 11
AAB69499
ID AAB69499 standard; protein; 111 AA.
XX
AC AAB69499;
XX
DT 23-APR-2001 (first entry)
XX
DE Rhizobium leguminosarum PII protein.
XX
KW Rhizobium leguminosarum; PII; plant nitrogen regulatory gene; P-PII;
KW nitrogen assimilation; transgenic plant; herbicide screening.
XX
OS Rhizobium leguminosarum.
XX
PN US6177275-B1.
XX
PD 23-JAN-2001.
XX
PF 23-JUL-1997; 97US-00899330.
XX

CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 160 AA;

Query Match 44.1%; Score 245.5; DB 4; Length 160;
Best Local Similarity 44.0%; Pred. No. 1.1e-20;
Matches 48; Conservative 30; Mismatches 30; Indels 1; Gaps 1;

QY 4 VEAIRPWRVSVQSSALLKIGIRGTVSDVRGFGAQQGSTERQGSFSEDKFVAKVOME 63
Db 52 VTVIKPKFLEDVREALSIGIQGLTVTEVKGFGRQKGHAELYRGAEYSVN-FLPKVKID 110

QY 64 IVVSKDOVEDVIEKIIIEARTGEIGDKIFLLPVSDVIRVRTGERGXA 112
Db 111 VAIADDQLDEVIDIVSKAAYTGKIGDKGKIFVAELQVRIRRTGEADEAA 159

RESULT 14
AAB79181
ID AAB79181 standard; protein; 89 AA.
AC AAB79181;
XX
DT 30-APR-2001 (first entry)
DE
XX Corynebacterium glutamicum HA protein sequence SEQ ID NO:318.

XX Corynebacterium glutamicum; homeostasis; adaptation; HA protein;
KW fine chemical production; organic acid; proteinogenic amino acid;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW carbohydrate; aromatic compound; vitamin; cofactor; polypeptide; enzyme;
KW diagnosis; Corynebacterium diphtheriae; genetic engineering;
KW Brevibacterium; environmental condition.
XX
OS Corynebacterium glutamicum.
XX
XX WO200100842-A2.
XX
PD 04-JAN-2001.
XX
XX 23-JUN-2000; 2000WO-IB000911.
XX
XX 25-JUN-1999; 99US-0141031P.
PR 08-JUL-1999; 99DE-01031636.
PR 09-JUL-1999; 99DE-01032125.
PR 09-JUL-1999; 99DE-01032126.
PR 09-JUL-1999; 99DE-01032127.
PR 09-JUL-1999; 99DE-01032128.
PR 09-JUL-1999; 99DE-01032129.
PR 09-JUL-1999; 99DE-01032226.
PR 14-JUL-1999; 99DE-01032920.
PR 14-JUL-1999; 99DE-01032922.
PR 14-JUL-1999; 99DE-01032924.
PR 14-JUL-1999; 99DE-01032928.
PR 14-JUL-1999; 99DE-01032930.
PR 14-JUL-1999; 99DE-01032933.
PR 14-JUL-1999; 99DE-01032935.
PR 14-JUL-1999; 99DE-01032973.
PR 14-JUL-1999; 99DE-01033002.
PR 14-JUL-1999; 99DE-01033003.
PR 14-JUL-1999; 99DE-01033005.

PR 14-JUL-1999; 99DE-01033006.
PR 31-AUG-1999; 99DE-01041378.
PR 31-AUG-1999; 99DE-01041379.
PR 31-AUG-1999; 99DE-01041390.
PR 31-AUG-1999; 99DE-01041391.
PR 03-SEP-1999; 99DE-01042088.
XX (BADI) BASF AG.
PA Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
PI WPI; 2001-061974/07.
XX N-PSDB; AAF71296.
DR
DR
XX
PT New isolated Corynebacterium glutamicum nucleic acid for production or
PT modulation of production of fine chemicals such as amino acids,
PT nucleosides, nucleotides, lipids, fatty acids, carbohydrates, vitamins or
PT enzymes.
XX
PS Claim 20; Page 563; 712pp; English.
XX
CC AAF71138 to AAF71137 encode the Corynebacterium glutamicum homeostasis
CC and adaptation (HA) proteins given in AAB79023 to AAB79242. The C.
CC glutamicum HA genes (I) can be used in vectors for expression in host
CC cells and production of fine chemicals, such as, an organic acid,
CC proteinogenic or nonproteinogenic amino acid (preferred), purine or
CC pyrimidine base, nucleoside, nucleotide, lipid, saturated or unsaturated
CC fatty acid, diol, carbohydrate, aromatic compound, vitamin, cofactor,
CC polypeptide or enzyme. The amino acids produced can be lysine, glutamine,
CC glutamate, alanine, aspartate, glycine, serine, threonine, methionine,
CC cysteine, valine, leucine, isoleucine, arginine, proline, histidine,
CC tyrosine, phenylalanine, or tryptophan. The fine chemical production can
CC be modulated. The presence of (I) or HA proteins encoded by then are used
CC for diagnosing the presence or activity of Corynebacterium diphtheriae.
CC (I) can be used to map the C. glutamicum genome or can be used as markers
CC for genetically engineered Corynebacterium or Brevibacterium. The HA
CC proteins encoded by the (I) are used to maintain homeostasis in C.
CC glutamicum or help the microorganism to adapt to different environmental
CC conditions
XX
SQ Sequence 89 AA;

Query Match 43.2%; Score 240.5; DB 4; Length 89;
Best Local Similarity 49.4%; Pred. No. 2e-20;
Matches 44; Conservative 25; Mismatches 19; Indels 1; Gaps 1;

QY 24 GIRGTVSDVRGFGAQQGSTERQGSFSEDKFVAKVMEIVVSKDOVEDVIEKIIIEAR 83
Db 1 GVQGMTVTETQGFQCKGHTVEYRGAEYAVD-FVPKVIIEVVISDAQAEVINIIVETAR 59

QY 84 TGEIGDGKIFLLPVSDVIRVRTGERGDKA 112
Db 60 TSKVGDGKVVMTTIEELVIRVRTGERGEAA 88

RESULT 15
ABB54937
ID ABB54937 standard; protein; 112 AA.
XX
AC ABB54937;
XX
DT 29-AUG-2003 (revised)
DT 16-MAY-2002 (first entry)
XX
DE Lactococcus lactis protein glnB.
XX
KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
XX
OS Lactococcus lactis; IL1403.
XX
PN FR2807446-A1.
XX
PD 12-OCT-2001.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 19:10:08 ; Search time 14 Seconds
(without alignments)
776.403 Million cell updates/sec

Title: US-09-756-541-2
Perfect score: 557
Sequence: 1 FYKVEAILRPWRVSVSSAL.....LLPVSVDVIRVTGERGDKAE 113

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	509	91.4	196	2	D85024	P II nitrogen sens
2	315.5	56.6	112	2	AH2095	nitrogen regulator
3	314.5	56.5	112	2	F70310	nitrogen regulator
4	303.5	54.5	112	2	H81961	nitrogen regulator
5	302.5	54.3	112	2	B81019	nitrogen regulator
6	300.5	53.9	112	2	A39696	nitrogen regulator
7	281.5	50.5	112	2	S13078	nitrogen regulator
8	278.5	50.0	112	2	S73175	nitrogen regulator
9	278.5	50.0	112	2	AF0354	nitrogen regulator
10	277.5	49.8	114	2	D82102	nitrogen regulator
11	276.5	49.6	112	2	S04377	nitrogen regulator
12	274.5	49.3	112	1	RGECp2	nitrogen regulator
13	274.5	49.3	112	2	C91056	hypothetical prote
14	274.5	49.3	112	2	G85900	hypothetical prote
15	274.5	49.3	112	2	AH0826	nitrogen regulator
16	270.5	48.6	112	2	S52328	nitrogen regulator
17	269.5	48.4	112	2	S33180	nitrogen regulator
18	268.5	48.2	112	2	F64062	nitrogen regulator
19	268.5	48.2	112	2	AB2794	nitrogen regulator
20	268.5	48.2	157	2	A97573	nitrogen regulator
21	267.5	48.0	112	2	C87493	nitrogen regulator
22	266.5	47.8	112	2	AD3374	nitrogen regulator
23	265.5	47.7	121	2	B82302	nitrogen regulator
24	264.5	47.5	112	2	C64307	probable nitrogen
25	260.5	46.8	112	2	G64467	nitrogen regulator
26	260.5	46.8	112	2	G70747	probable nitrogen
27	253.5	45.5	112	2	D82985	nitrogen regulator
28	251.5	45.2	112	2	B69468	nitrogen regulator
29	251.5	45.2	112	2	T35668	nitrogen regulator

30	251	45.1	111	2	B26567	nitrogen regulator
31	249.5	44.8	115	2	D69188	nitrogen regulator
32	248.5	44.6	112	2	H82631	nitrogen regulator
33	248.5	44.6	112	2	AB0560	nitrogen regulator
34	248	44.5	111	2	A33600	nitrogen regulator
35	246.5	44.3	85	2	S76404	nitrogen regulator
36	246.5	44.3	112	2	AG0381	nitrogen regulator
37	245.5	44.1	112	2	B64775	nitrogen regulator
38	245.5	44.1	112	2	H90691	nitrogen regulator
39	245.5	44.1	112	2	D85542	nitrogen regulator
40	244.5	43.9	116	2	AD2915	nitrogen regulator
41	244.5	43.9	116	2	G97689	glnK protein (AJ00
42	243.5	43.7	115	2	B69188	nitrogen regulator
43	237.5	42.6	112	2	D87415	nitrogen regulator
44	234.5	42.1	114	2	H87313	nitrogen regulator
45	218	39.1	113	2	G86823	nitrogen regulator

ALIGNMENTS

RESULT 1

D85024
P II nitrogen sensing protein GLB I [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cross)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C;Accession: D85024
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: D85024
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-196 <STO>
A;Cross-references: GB:NC_001268; NID:g7268574; PIDN:CAB80683.1; GSPDB:GN00140
C;Genetics:
A;Gene: AR4901900
A;Map position: 4

Query Match 91.4%; Score 509; DB 2; Length 196;
Best Local Similarity 90.3%; Pred. No. 6.2e-40;
Matches 102; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY	1	FYKVEAILRPWRVSVSSALLKIGIRGVTVSDVRGFGAQQGGSTERQGGSEFSEDKFVAKV	60
DB	74	FYKVEAILRPWRVSVSSALLKIGIRGVTVSDVRGFGAQQGGSTERHGGSEFSEDKFVAKV	133
QY	61	KMEIVWSKDQVEDVIEKIEEARTGEIGDGKIFLLPVSVDVIRVTGERGDKAE	113
DB	134	KMEIVVKDQVESVINTIEGARTGEIGDGKIFVLPSVDVIRVTGERGEKAE	186

RESULT 2

AH2095
nitrogen regulatory protein P-II glnB [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AH2095
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH2095
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-112 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB74018.1; PID:g17131411; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: glnB

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 19:06:17 ; Search time 10.5 Seconds
(without alignments)
560.374 Million cell updates/sec

Title: US-09-756-541-2
Perfect score: 557
Sequence: 1 FYKVEAILRPWRVSVSSAL.....LLPVSDVIRVTRGGRGKAE 113

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	315.5	56.6	112	1	GLNB_NOSPU
2	314.5	56.5	112	1	GLNB_AQUAE
3	314.5	56.5	112	1	GLNB_FREDI
4	311.5	55.9	112	1	GLNB_SYNY3
5	306.5	55.0	112	1	GLNB_SYNP7
6	286.5	51.4	112	1	GLNB_PASMP
7	281.5	50.5	112	1	GLNB_AZOB
8	278.5	50.0	112	1	GLNB_PORPU
9	276.5	49.6	112	1	GLNB_KLEOX
10	274.5	49.3	112	1	GLNB_ECOLI
11	270.5	48.6	112	1	GLNB_RHORI
12	269.5	48.4	112	1	GLNB_RHIME
13	269.5	48.4	112	1	GLNB_RHOSH
14	268.5	48.2	112	1	GLNB_HAEIN
15	267.5	48.0	112	1	GLNB_RHILO
16	264.5	47.5	112	1	Y059_METJA
17	260.5	46.8	112	1	GLNB_BRAJA
18	260.5	46.8	112	1	GLNB_MYCTU
19	260.5	46.8	112	1	YD44_METJA
20	255.5	45.9	112	1	GLNB_RHOCA
21	254.5	45.7	113	1	GLNB_CYACA
22	251	45.1	111	1	GLNB_RHILV
23	249.5	44.8	115	1	GLN2_METTH
24	245.5	44.1	112	1	GLNK_ECOLI
25	243.5	43.7	115	1	GLN1_METTH
26	232.5	41.7	112	1	GLNB_RHIET
27	195.5	35.1	116	1	NRGE_BACSU
28	181.5	32.6	123	1	GLN2_METBA
29	170.5	30.6	125	1	GLN4_METBA
30	167.5	30.1	128	1	GLN2_METTL
31	166	29.8	121	1	GLN2_METMP
32	165	29.6	122	1	GLN2_METIV
33	163.5	29.4	105	1	GLN1_METIV

34	159	28.5	121	1	GLN2_METTM	Q50787 methanobact
35	153.5	27.6	105	1	GLN1_METTL	P25771 methanococc
36	149.5	26.8	105	1	GLN1_METTM	Q50786 methanobact
37	144.5	25.9	105	1	GLN1_METBA	P54808 methanosarc
38	140.5	25.2	105	1	GLN1_METMP	P71524 methanococc
39	135.5	24.3	105	1	GLN3_METBA	P54807 methanosarc
40	79.5	14.3	307	1	GSHB_PROMP	Q77ug9 prochloroco
41	75.5	13.6	305	1	FAR1_HELAS	Q07981 helix asper
42	75	13.5	187	1	KIP2_MOUSE	Q92309 mus musculu
43	72.5	13.0	529	1	IMDH_MYCTU	Q50715 mycobacteri
44	72	12.9	187	1	KIP2_HUMAN	O75838 homo sapien
45	70.5	12.7	573	1	CH60_HUMAN	P10809 homo sapien

ALIGNMENTS

RESULT 1
GLNB_NOSPU
ID GLNB_NOSPU STANDARD; PRT; 112 AA.
AC O30794;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Nitrogen regulatory protein P-II (PII signal transducing protein).
GM GLNB.
OS Nostoc punctiforme.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=63737;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29133 / PCC 73102;
RX MEDLINE=98304077; PubMed=9639924;
RA Hanson T.E., Forchhammer K., Tandeau de Marsac N., Meeks J.C.;
RT *Characterization of the glnB gene product of Nostoc punctiforme
RT strain ATCC 29133: glnB or the PII protein may be essential.";
RL Microbiology 144:1537-1547(1998).
CC -!- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE GS
CC GENE (GLNA). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO
CC NR-I-PHOSPHATE, THE TRANSCRIPTIONAL ACTIVATOR OF GLNA. WHEN P-II
CC IS PHOSPHORYLATED, THESE EVENTS ARE REVERSED (BY SIMILARITY).
CC -!- FUNCTION: IN NITROGEN-LIMITING CONDITIONS, WHEN THE RATIO OF GLN
CC TO 2-KETOGLUTARATE DECREASES, P-II IS PHOSPHORYLATED WHICH ALLOWS
CC THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS), THUS ACTIVATING
CC THE ENZYME (BY SIMILARITY).
CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- PTM: PHOSPHORYLATION DEPENDENT ON THE NITROGEN SOURCE AND SPECTRAL
CC LIGHT QUALITY (BY SIMILARITY).
CC -!- SIMILARITY: Belongs to the P(II) protein family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement [See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch].

EMBL; AF017419; AAC26348.1; -
HSSP; P05826; 2PII.
InterPro; IPR002187; PII_glnB.
InterPro; IPR002332; PII_glnB_UMP_s.
Pfam; PF00543; P-II; 1.
PRINTS; PR00340; PIIGLNB.
ProDom; PD001194; PII_glnB; 1.
PROSITE; PS00496; PII_GLNB_UMP; 1.
PROSITE; PS00638; PII_GLNB_CTER; 1.
KW Transcription regulation; Nitrogen fixation; Phosphorylation.
FT MOD RES 49 49 PHOSPHORYLATION (PROBABLE).
SQ SEQUENCE 112 AA; 12478 MW; 9C2224C38B67583A CRC64;

Query Match 56.6%; Score 315.5; DB 1; Length 112;
Best Local Similarity 58.2%; Pred. No. 1e-22;

Matches	64;	Conservative	22;	Mismatches	23;	Indels	1;	Gaps	1;
QY	3	KVEAILRPWRVSVSSALLKIGIRGVTVDVRGFGAQQGSTERQGGSEFSEDKFVAKVKM	62						
Db	3	KVEAILRPFKLDEVKIALVNAGIVGMTVSEVRGFGQRKGQTERYRGSEYTVF-FLQKLKV	61						
QY	63	EIVVSKDQVEDVIEKIIIEARTGEIGDKIFLLPVSDVIRVTRTGERGDKA	112						
Db	62	EIVVDNQVDMVVDKIIAARTGEIGDKIFISPVQEVIRIRTKENKTEA	111						
RESULT 2									
GLNB_AQUAE		STANDARD;		PRT;		112 AA.			
ID	GLNB_AQUAE								
AC	O66513;								
DT	30-MAY-2000 (Rel. 39, Created)								
DT	30-MAY-2000 (Rel. 39, Last sequence update)								
DT	16-OCT-2001 (Rel. 40, Last annotation update)								
DE	Nitrogen regulatory protein P-II.								
GN	GLNB OR AQ 109.								
OS	Aquifex aeolicus.								
OC	Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.								
OX	NCBI_TaxID=63363;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=VFS;								
RX	MEDLINE=98196666; PubMed=9537320;								
RA	Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,								
RA	Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,								
RA	Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;								
RT	"The complete genome of the hyperthermophilic bacterium Aquifex								
RT	aeolicus.";								
RL	Nature 392:353-358(1998).								
CC	-!- FUNCTION: IN NITROGEN-LIMITING CONDITIONS, WHEN THE RATIO OF GLN								
CC	TO 2-KETOGLUTARATE DECREASES, P-II IS URIDYLATED TO P-II-UMP. P-								
CC	II-UMP ALLOWS THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS),								
CC	THUS ACTIVATING THE ENZYME. CONVERSELY, IN NITROGEN EXCESS P-II								
CC	IS DEURIDYLATED AND PROMOTES THE ADENYLATION OF GS (BY								
CC	SIMILARITY).								
CC	-!- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE GS								
CC	GENE (GLNA). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO								
CC	NR-I-PHOSPHATE, THE TRANSCRIPTIONAL ACTIVATOR OF GLNA. WHEN P-II								
CC	IS URIDYLATED TO P-II-UMP, THESE EVENTS ARE REVERSED (BY								
CC	SIMILARITY).								
CC	-!- SUBUNIT: Homotrimer (By similarity).								
CC	-!- SIMILARITY: Belongs to the P(II) protein family.								

This SWISS-PROT entry is copyright. It is produced through a collaboration									
between the Swiss Institute of Bioinformatics and the EMBL outstation -									
the European Bioinformatics Institute. There are no restrictions on its									
use by non-profit institutions as long as its content is in no way									
modified and this statement is not removed. Usage by and for commercial									
entities requires a license agreement (See http://www.isb-sib.ch/announce/									
or send an email to license@isb-sib.ch).									

DR	EMBL; AE000674; AAC06473.1; -.								
DR	PIR; F70310; F70310.								
DR	HSSP; P05826; 2PII.								
DR	InterPro; IPR002187; PII_glnB.								
DR	InterPro; IPR002332; PII_GlnB_UMP_S.								
DR	Pfam; PF00543; P-II; 1.								
DR	PRINTS; PR00340; PIIGLNB.								
DR	ProDom; PD001194; PII_glnB; 1.								
DR	PROSITE; PS00496; PII_GlnB_UMP; FALSE_NEG.								
DR	PROSITE; PS00638; PII_GlnB_CTER; 1.								
KW	Transcription regulation; Nitrogen fixation; Complete proteome.								
FT	BINDING 51 51 UMP (BY SIMILARITY).								
SQ	SEQUENCE 112 AA; 12497 MW; 0E44B4B171A6233B CRC64;								
Query Match									
Best Local Similarity		56.5%;		Score 314.5;		DB 1;		Length 112;	
Matches		59;		Conservative		29;		Mismatches 21; Indels 1; Gaps 1;	

RESULT 4
Glnb_Synp3
ID Glnb_Synp3 STANDARD; PRT; 112 AA.
AC Q55247;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nitrogen regulatory protein P-II (PII signal transducing protein).
GN Glnb OR SSI0707.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98088000; PubMed=9426594;
RA Garcia-Dominguez M., Florencio F.J.;
RT "Nitrogen availability and electron transport control the expression
RT of glnB gene (encoding PII protein) in the cyanobacterium
RT Synechocystis sp. PCC 6803.";
RL Plant Mol. Biol. 35:723-734(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirose M., Sugita M., Sasanoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -!- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE GS
CC GENE (GLN). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO
CC NR-I-PHOSPHATE, THE TRANSCRIPTIONAL ACTIVATOR OF GLN. WHEN P-II
CC IS PHOSPHORYLATED, THESE EVENTS ARE REVERSED (BY SIMILARITY).
CC -!- FUNCTION: IN NITROGEN-LIMITING CONDITIONS, WHEN THE RATIO OF GLN
CC TO 2-KETOGLUTARATE DECREASES, P-II IS PHOSPHORYLATED WHICH ALLOWS
CC THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS), THUS ACTIVATING
CC THE ENZYME (BY SIMILARITY).
CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- PTM: PHOSPHORYLATION DEPENDENT ON THE NITROGEN SOURCE AND SPECTRAL
CC LIGHT QUALITY (BY SIMILARITY).
CC -!- SIMILARITY: Belongs to the P(II) protein family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X97496; CAA66127.1; -;
DR EMBL; D90915; BAA18533.1; ALT_INIT.
DR HSSP; P05826; 2PII.
DR InterPro; IPR002187; PII_glnB.
DR InterPro; IPR002332; PII_Glnb_UMP_S.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR03340; PIIGLNB.
DR ProDom; PD031194; PII_glnB; 1.
DR PROSITE; PS00496; PII_Glnb_UMP; 1.
DR PROSITE; PS00638; PII_Glnb_CTER; 1.
KW Transcription regulation; Nitrogen fixation; phosphorylation;
KW Complete proteome.
FT MOD_RES 49 49 PHOSPHORYLATION (PROBABLE).
SQ SEQUENCE 112 AA; 12397 MW; F9ABD0F5C173B799 CRC64;

Query Match 55.9%; Score 311.5; DB 1; Length 112;
Best Local Similarity 55.5%; Pred. No. 2.5e-22;
Matches 61; Conservative 25; Mismatches 23; Indels 1; Gaps 1;

3 KVEAILRPWRVSVSSALLKIGIRGVTVSDVRGFGAQQGSTERQGGSEFSEDKFVAKVM 62

Db 3 KVEAILRPFKLDEVKIALVNAVIGMTVSEVRGFGKQGTERTYRGSEYTVF-FLOKLKI 61
Qy 63 EIVVSKDQVEDVIEKIEEARTGEIGDKIFLPLVSDVIRVTGERGDKA 112
Db 62 EIVVDEGVDMVVDKLVSAARTGEIGDKIFISFVDSVVRIRTKGKTEA 111

RESULT 5
Glnb_Synp7
ID Glnb_Synp7 STANDARD; PRT; 112 AA.
AC P80016;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nitrogen regulatory protein P-II (PII signal transducing protein).
GN Glnb.
OS Synechococcus sp. (strain PCC 7942) (Anacystis nidulans R2), and
OS Synechococcus sp. (strain PCC 6301) (Anacystis nidulans).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=1140, 1139;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9271233; PubMed=1905010;
RA Tsinoenmas N.F., Castets A.M., Harrison M.A., Allen J.F.,
RA Tandeau de Marsac N.;
RT "Photosynthetic electron transport controls nitrogen assimilation in
RT cyanobacteria by means of posttranslational modification of the glnB
RT gene product.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:4565-4569(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=PCC 6301;
RA Inoue K., Bryant D.A.;
RT "Genes required for c-type cytochrome biogenesis.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP SEQUENCE OF 1-30.
RX MEDLINE=90249505; PubMed=2110911;
RA Harrison M.A., Keen J.N., Findlay J.B.C., Allen J.F.;
RT "Modification of a glnB-like gene product by photosynthetic electron
RT transport in the cyanobacterium Synechococcus 6301.";
RL FEBS Lett. 264:25-28(1990).
RN [4]
RP SEQUENCE OF 1-40.
RX STRAIN=PCC 6301;
MEDLINE=91355213; PubMed=1653017;
RA Li N., Warren P.V., Golbeck J.H., Frank G., Zuber H., Bryant D.A.;
RT "Polypeptide composition of the Photosystem I complex and the
RT Photosystem I core protein from Synechococcus sp. PCC 6301.";
RL Biochim. Biophys. Acta 1059:215-225(1991).
RN [5]
RP PHOSPHORYLATION.
RX STRAIN=PCC 7942;
MEDLINE=94110245; PubMed=8282715;
RA Forchhammer K., Tandeau de Marsac N.;
RT "The PII protein in the cyanobacterium Synechococcus sp. strain PCC
RT 7942 is modified by serine phosphorylation and signals the cellular
RT N-status.";
RL J. Bacteriol. 176:84-91(1994).
RN [6]
RP PHOSPHORYLATION.
RX STRAIN=PCC 7942;
MEDLINE=96011366; PubMed=7592328;
RA Forchhammer K., Tandeau de Marsac N.;
RT "Phosphorylation of the PII protein (glnB gene product) in the
RT cyanobacterium Synechococcus sp. strain PCC 7942: analysis of in
RT vitro kinase activity.";
RL J. Bacteriol. 177:5812-5817(1995).
CC -!- FUNCTION: P-II INDIRECTLY PREVENTS THE TRANSCRIPTION OF THE GS
CC GENE (GLN). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO

```
CC NR-1-PHOSPHATE, THE TRANSCRIPTIONAL ACTIVATOR OF GLNA. WHEN P-II
CC IS PHOSPHORYLATED, THESE EVENTS ARE REVERSED.
CC
CC -!- FUNCTION: IN NITROGEN-LIMITING CONDITIONS, WHEN THE RATIO OF GLN
CC TO 2-KETOGLUTARATE DECREASES, P-II IS PHOSPHORYLATED WHICH ALLOWS
CC THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS), THUS ACTIVATING
CC THE ENZYME.
CC
CC -!- SUBUNIT: Homotrimer.
CC
CC -!- PTM: PHOSPHORYLATION DEPENDENT ON THE NITROGEN SOURCE AND SPECTRAL
CC LIGHT QUALITY.
CC
CC -!- SIMILARITY: Belongs to the P(II) protein family.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M62447; AAA27312.1; -.
CC EMBL; AF079137; AAF04333.1; -.
CC PIR; A39696; A39696.
CC HSSP; P05826; 2PII.
CC PhosSite; P80016; -.
CC InterPro; IPR002187; PII_glnB.
CC InterPro; IPR002332; PII_GlnB_UMP_S.
CC Pfam; PF00543; P-II; 1.
CC PRINTS; PR00340; PIIGLNB.
CC ProDom; PD001194; PII_glnB; 1.
CC PROSITE; PS00496; PII_GlnB_UMP; 1.
CC PROSITE; PS00638; PII_GlnB_CTER; 1.
CC Transcription regulation; Nitrogen fixation; Phosphorylation.
KW MOD RES 49 49 PHOSPHORYLATION.
FT CONFLICT 83 83 T -> P (IN REF. 1; AAA27312).
FT SEQUENCE 112 AA; 12391 MW; 5F44B64CBFF3C559 CRC64;
SQ
Query Match 55.0%; Score 306.5; DB 1; Length 112;
Best Local Similarity 56.4%; Pred. No. 7.2e-22;
Matches 62; Conservative 21; Mismatches 26; Indels 1; Gaps 1;
QY 3 KVEAILRPWRVSVSSALLKIGIRGVTVSDVRGFGAQQGSTERQGSGSEFSEDKFVAKVKM 62
DB 3 KIEAIRPFKLDVXIALVNAVIGVMTVSEVRGFGKQGTERTYRGSEYTVF-FLQKLKL 61
QY 63 EIVVSKDQVEDVIEKIIIEARTGEIGDKIFLLPVSVDVIRVTGERGDKA 112
DB 62 EIVVEDAQVTDVTDKIVAAARTGEIGDKIFVSPVDQIRIRTKGNADA 111
RESULT 6
GLNB_PASMU STANDARD; PRT; 112 AA.
AC Q9CJL1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nitrogen regulatory protein P-II.
GN GLNB OR PM2004.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465 (2001).
CC -!- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE
CC GLUTAMINE SYNTHETASE GENE (GLNA). P-II PREVENTS NR-II CATALYZED
CC CONVERSION OF NR-I TO NR-I-PHOSPHATE, THE TRANSCRIPTIONAL
CC ACTIVATOR OF GLNA. WHEN P-II IS URIDYLYLATED TO P-II-UMP, THESE
```

```
CC EVENTS ARE REVERSED. WHEN THE RATIO OF GLN TO 2-KETOGLUTARATE
CC DECREASES, P-II IS URIDYLYLATED TO P-II-UMP, WHICH CAUSES THE
CC DEADENYLATION OF GLUTAMINE SYNTHETASE BY GLNE, SO ACTIVATING THE
CC ENZYME (BY SIMILARITY).
CC
CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- PTM: Uridylylated/deuridylylated by glnD (By similarity).
CC -!- SIMILARITY: Belongs to the P(II) protein family.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AE006236; AAK04088.1; -.
CC HSSP; P05826; 2PII.
CC InterPro; IPR002187; PII_glnB.
CC InterPro; IPR002332; PII_GlnB_UMP_S.
CC Pfam; PF00543; P-II; 1.
CC PRINTS; PR00340; PIIGLNB.
CC ProDom; PD001194; PII_glnB; 1.
CC PROSITE; PS00496; PII_GlnB_UMP; 1.
CC PROSITE; PS00638; PII_GlnB_CTER; 1.
CC Transcription regulation; Nitrogen fixation; Complete proteome.
KW BINDING 51 51 UMP (BY SIMILARITY).
FT SEQUENCE 112 AA; 12684 MW; 0CC9D24651F47642 CRC64;
SQ
Query Match 51.4%; Score 286.5; DB 1; Length 112;
Best Local Similarity 52.7%; Pred. No. 5.2e-20;
Matches 58; Conservative 23; Mismatches 28; Indels 1; Gaps 1;
QY 3 KVEAILRPWRVSVSSALLKIGIRGVTVSDVRGFGAQQGSTERQGSGSEFSEDKFVAKVKM 62
DB 3 KIEAIRPFKLDVRESLSDVGITGMTVTEVRGFGKQGTERTYRGAEYMWVD-FLPKVKM 61
QY 63 EIVVSKDQVEDVIEKIIIEARTGEIGDKIFLLPVSVDVIRVTGERGDKA 112
DB 62 EIVTDEQVDCIEAIMEAQTGKIGDKIFVDVIRVIRIRTKGEENEDA 111
RESULT 7
GLNB_AZOBR STANDARD; PRT; 112 AA.
AC P21193;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Nitrogen regulatory protein P-II.
GN GLNB.
OS Azospirillum brasilense.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Rhodospirillaceae; Azospirillum.
OX NCBI_TaxID=192;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sp7;
RX MEDLINE=91094780; PubMed=1702507;
RA de Zamaroczy M., Delorme F., Elmerich C.;
RT "Characterization of three different nitrogen-regulated promoter
RT regions for the expression of glnB and glnA in Azospirillum
RT brasilense."
RL Mol. Gen. Genet. 224:421-430 (1999).
RN [2]
RP SEQUENCE OF 88-112 FROM N.A.
RC STRAIN=Sp7;
RX MEDLINE=87076765; PubMed=2878685;
RA Bozouklian H., Elmerich C.;
RT "Nucleotide sequence of the Azospirillum brasilense Sp7 glutamine
RT synthetase structural gene."
RL Biochimie 68:1181-1187 (1986).
CC -!- FUNCTION: IN NITROGEN-LIMITING CONDITIONS, WHEN THE RATIO OF GLN
```

between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

Query Match 50.5%; Score 281.5; DB 1; Length 112;
Best Local Similarity 49.1%; Pred. No. 1.5e-19;
Matches 54: Conservative 26; Mismatches 29; Indels 1; Gaps 1;

QY	63	EIVSKQVEDVIEKIIIEEARTGEIGDKIFLLPVSDVIRVRTGERGDKA	112
		: : : : : : : : : : : : : :	
AK	62	EIVMEDS:MERATEAIOCAAAHTGRIGDKIFVTPVEEVRIRTEGKGDA	111

CC -i- SIMILARITY: Belongs to the P(II) protein family.
CC -----
CC


```

by protein phosphorylation." ;
RL Biochimie 71:1005-1012(1989) .
RN [17]
RP CRYSTALLIZATION, AND SUBUNITS.
RC SPECIES=E.coli;
RX MEDLINE=94123764; PubMed=8293810;
RA Vasudevan S.G., Gedye C., Dixon N.E., Cheah E., Carr P.D.,
RA Suffolk P.M., Jeffrey P.D., Ollis D.L.;
RA "Escherichia coli pII protein: purification, crystallization and
RT oligomeric structure." ;
RL FEBS Lett. 337:255-258(1994) .
RN [18]
RP X-RAY CRYSTALLOGRAPHY {2.7 ANGSTROMS}.
RC SPECIES=E.coli;
RX MEDLINE=95171116; PubMed=7866749;
RA Cheah E., Carr P.D., Suffolk P.M., Vasuvedan S.G., Dixon N.E.,
RA Ollis D.L.;
RA "Structure of the Escherichia coli signal transducing protein PII." ;
RT "Structure 2:981-990(1994)".
RL

Query Match 49.3%; Score 274.5; DB 1; Length 112;
Best Local Similarity 50.0%; Pred. No. 6.8e-19;
Matches 55; Conservative 25; Mismatches 29; Indels 1; Gaps 1;

QY 3 KVEAILRPWRVSVSSALLKIGIRGVTVDVRGFGAQQGSTERQGGSEFSEDKFVAKVKM 62
DB 3 KIDAIIPFKLDDVREALAEVGITGMTVTEVKGFGKRGKHTELRGAEYMDV-FLPKVKI 61
QY 63 EIVVSKDQVEDVIEKIEEARTGEIGDGKIFLLPVSVDVIRVTGERGDKA 112
DB 62 EIVVPEIDVDTCTVIRTACTGKIGDKGKIFVEDVARVIRRTGEEDAA 111

RESULT 11
GLNB_RHORU STANDARD; PRT; 112 AA.
AC Q53044;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Nitrogen regulatory protein P-II.
GN GLNB.
OS Rhodospirillum rubrum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Rhodospirillaceae; Rhodospirillum.
OX NCBI_TaxID=1085;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 11170 / S1;
RX MEDLINE=96254013; PubMed=8704966;
RA Johansson M., Nordlund S.;
RT "Transcription of the glnB and glnA genes in the photosynthetic
RT bacterium Rhodospirillum rubrum." ;
RL Microbiology 142:1265-1272(1996) .
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 11170 / S1;
RA Zhang Y., Ludden P.W., Roberts G.P.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE
CC GLUTAMINE SYNTHETASE GENE (GLNA). P-II PREVENTS NR-II CATALYZED
CC CONVERSION OF NR-I TO NR-I-PHOSPHATE, THE TRANSCRIPTIONAL
CC ACTIVATOR OF GLNA. WHEN P-II IS URIDYLATED TO P-II-UMP, THESE
CC EVENTS ARE REVERSED. WHEN THE RATIO OF GLN TO 2-KETOGLUTARATE
CC DECREASES, P-II IS URIDYLATED TO P-II-UMP, WHICH CAUSES THE
CC DEADENYLATION OF GLUTAMINE SYNTHETASE, SO ACTIVATING THE ENZYME.
CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- SIMILARITY: Belongs to the P(II) protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```


CC -!- SIMILARITY: Belongs to the P(II) protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC EMBL; AP002994; BAB47945.1; -.
DR InterPro; IPR002187; PII_glnB.
DR InterPro; IPR002332; PII_glnB_UMP_S.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR00340; PIIGLNB.
DR ProDom; PD001194; PII_glnB; 1.
DR PROSITE; PS00496; PII_GlnB_UMP; 1.
DR PROSITE; PS00638; PII_GlnB_CTER; 1.
KW Transcription regulation; Nitrogen fixation; Complete proteome.
FT BINDING 51 UMP (BY SIMILARITY).
SQ SEQUENCE 112 AA; 12200 MW; CPQF60FE7185465A CRC64;

Query Match 48.0%; Score 267.5; DB 1; Length 112;
Best Local Similarity 47.7%; Pred. No. 3.1e-18;
Matches 51; Conservative 27; Mismatches 28; Indels 1; Gaps 1;

Qy 3 KVEAILRPWRVSVSSALLKIGIRGVTVSDVRGFGAQQGSTERQGGSEFSEDKFVAKVKM 62
Db 3 KIEAILKPKLDEVKEALQEAQLQGITVTEAKGFGKQKHTLYRGAIEYVVD-FLPKVKI 61
Qy 63 EIVVSKDQVEDVIEKIIIEARTGEIGDKIPLLPVSDVIRVTRTGERG 109
Db 62 EVVLGDDAVEGAIEAIRKAAQTGRIGDGKIFVSNIEEVVRIRTGTG 108

Search completed: May 24, 2004, 19:13:35
Job time : 11.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: May 24, 2004, 19:09:38 ; Search time 36 Seconds
(without alignments)
990.377 Million cell updates/sec

Title: US-09-756-541-2
Perfect score: 557
Sequence: 1 FYKVEAILRPWRVSVSSAL.....LLPVSDVIRVTGERGDKAE 113

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 3.5

Searched: 1017041 seqs, 315518202 residues 1017041
Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL 25:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	557	100.0	197	10 Q9ZST5	Q9zst5 ricinus com
2	509	91.4	196	10 Q9ZST4	Q9zst4 arabidopsis
3	492	88.3	194	10 Q9ARI4	Q9ari4 medicago sa
4	315.5	56.6	112	16 Q9L422	Q9l422 anabaena sp
5	313.5	56.3	112	2 Q9LAQ3	Q9laq3 synechococc
6	311.5	55.9	112	16 Q8DLA5	Q8dla5 synechococc
7	303.5	54.5	112	16 Q9JWC4	Q9jwc4 neisseria m
8	302.5	54.3	112	16 Q9JXK6	Q9jxk6 neisseria m
9	302.5	54.3	112	16 Q7V5R4	Q7v5r4 prochloroco
10	295.5	53.1	112	2 Q8GQS4	Q8ggs4 acetobacter
11	295.5	53.1	112	16 Q7U8Z7	Q7u8z7 synechococc
12	292.5	52.5	112	16 Q7VA51	Q7va51 prochloroco
13	287.5	51.6	108	2 Q8KUJ2	Q8kuj2 neisseria m
14	286.5	51.4	112	2 Q9EZQ2	Q9ezq2 azoarcus sp
15	286.5	51.4	112	2 Q9L400	Q9l400 prochloroco
16	286.5	51.4	112	16 Q7V025	Q7v025 prochloroco

17	285.5	51.3	112	16 Q7VPK6	Q7vpk6 haemophilus
18	283.5	50.9	114	16 Q87MP6	Q87mf6 vibrio para
19	282.5	50.7	114	16 Q8DBE2	Q8dbe2 vibrio vuln
20	282.5	50.7	133	16 Q7UUZ1	Q7uuz1 rhodopirell
21	281.5	50.5	105	2 Q8KUJ5	Q8kuj5 neisseria m
22	279.5	50.2	112	16 Q8FP24	Q8fp24 corynebacte
23	279.5	50.2	112	16 Q8EBJ6	Q8ebj6 shewanella
24	278.5	50.0	112	2 Q31188	Q31188 rhodobacter
25	278.5	50.0	112	16 Q8ZCQ9	Q8zcq9 yersinia pe
26	277.5	49.8	114	16 Q9KFX3	Q9kpx3 vibrio chol
27	276.5	49.6	112	2 Q9AMM9	Q9amm9 azospirillu
28	275.5	49.5	112	16 Q7WJI4	Q7wji4 bordetella
29	275.5	49.5	112	16 Q7WAE1	Q7wae1 bordetella
30	275.5	49.5	112	16 Q7VVG9	Q7vvg9 bordetella
31	272.5	48.9	112	2 P70731	P70731 azospirillu
32	272.5	48.9	112	2 P94125	P94125 azorhizobiu
33	271.5	48.7	101	2 Q8KHS5	Q8khs5 neisseria m
34	270.5	48.6	101	2 Q8KUJ0	Q8kuj0 neisseria m
35	270.5	48.6	112	16 Q9X705	Q9x705 corynebacte
36	269.5	48.4	100	2 Q8KI77	Q8ki77 neisseria m
37	269.5	48.4	100	2 Q8KI78	Q8ki78 neisseria m
38	268.5	48.2	98	2 Q8KUJ4	Q8kuj4 neisseria m
39	268.5	48.2	99	2 Q8KUJ3	Q8kuj3 neisseria m
40	268.5	48.2	99	2 Q8KI58	Q8ki58 neisseria m
41	268.5	48.2	100	2 Q8KH45	Q8kh45 neisseria m
42	268.5	48.2	157	16 Q8UEI8	Q8uei8 agrobacteri
43	267.5	48.0	98	2 Q8KUI8	Q8kui8 neisseria m
44	267.5	48.0	99	2 Q8KHB9	Q8khb9 neisseria m
45	267.5	48.0	112	2 P94852	P94852 herbaspiril

ALIGNMENTS

RESULT 1

ID	Q9ZST5	PRELIMINARY:	PRT;	197 AA.
AC	Q9ZST5;			
DT	01-MAY-1999 (Tremblrel. 10, Created)			
DT	01-MAY-1999 (Tremblrel. 10, Last sequence update)			
DT	01-JUN-2003 (Tremblrel. 24, Last annotation update)			
DE	PII protein (Fragment).			
OS	Ricinus communis (Castor bean).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eurosids I; Malpighiales; Euphorbiaceae; Acalyphoideae; Acalypheae;			
OC	Ricinus.			
OX	NCBI_TaxID=3988;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99030678; PubMed=9811909;			
RA	Hsieh M.H., Lam H.M., van de Loo F.J., Coruzzi G.;			
RT	"A PII-like protein in Arabidopsis: putative role in nitrogen			
RT	sensing."			
RL	Proc. Natl. Acad. Sci. U.S.A. 95:13965-13970(1998).			
DR	EMBL; AF095454; AAC78332.1; --			
DR	HSSP; P05826; 2PII.			
DR	GO; GO:0030234; F:enzyme regulator activity; IEA.			
DR	GO; GO:0006808; P:regulation of nitrogen utilization; IEA.			
DR	InterPro; IPR002187; PII_glnB.			
DR	Pfam; PF00543; P-II; 1.			
DR	PRINTS; PR00340; PIIGLNB.			
DR	ProDom; PD001194; PII_glnB; 1.			
DR	PROSITE; PS00638; PII_GLNB_CTER; 1.			
FT	NCN TER			
SQ	SEQUENCE 197 AA; 21781 MW; 2D18A50F15406B20 CRC64;			

Query Match 100.0%; Score 557; DB 10; Length 197;
Best Local Similarity 100.0%; Pred. No. 2.3e-45;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYKVEAILRPWRVSVSSALLKIGIRGVTVDVRFGAQGGSTERQGGSEFSDKFKVAKV 60
|||||

Db 72 FYKVEAILRPWRVSVSSALLKIGIRGVTVDVRGFGAQQGSTERQGGSEFSEDKFVAKV 131

QY 61 KMEIVWSKDQVEDVIEKIIIEARTGEIGDGKIFLLPVSVDVIRVVTGERGDKAE 113

Db 132 KMEIVWSKDQVEDVIEKIIIEARTGEIGDGKIFLLPVSVDVIRVVTGERGDKAE 184

RESULT 2

Q9ZST4 PRELIMINARY; PRT; 196 AA.

AC Q9ZST4; 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE PII protein (P II nitrogen sensing protein GLB I) (At4g01900).

GN T7B11.16 OR AT4G01900.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99030678; PubMed=9811909;

RA Hsieh M.H., Lam H.M., van de Loo F.J., Coruzzi G.,

RT "A PII-like protein in Arabidopsis: putative role in nitrogen

RT sensing.";

RL Proc. Natl. Acad. Sci. U.S.A. 95:13965-13970(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RA Huang B.N., Nascimento L., de la Bastide M., Habermann K., Vil M.D.,

RA Preston R.R., Spiegel L.A., See L.H., Shah R., Matero A.,

RA O'Shaughnessy A., Rodriguez M., Shekher M., Swaby I., Schutz K.,

RA Parnell L.D., Dedhia N.N., McCombie W.R.;

RA "Arabidopsis thaliana BAC T7B11 from chromosome IV near 10 cM.";

RT Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Spiegel L.A., Huang B.N., Nascimento L.U., de la Bastide M., Vil D.M.,

RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,

RA Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,

RA Mewes H.W., Lemcke K., Mayer K.F.X.;

RA Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;

RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.

RN [5]

RP SEQUENCE FROM N.A.

RA Kim C.J., Chen H., Cheuk R., Shinn P., Bowser L., Carninci P.,

RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,

RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,

RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,

RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,

RA Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,

RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;

RT "Arabidopsis ORF clones.";

RL Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AF095455; AAC78333.1; -

DR EMBL; AC007138; AAD22652.1; -

DR EMBL; AL161493; CAB80683.1; -

DR EMBL; BT005209; AAO63273.1; -

DR PIR; D85024; D85024.

DR HSSP; P05826; 2PII.

DR GO; GO:0030234; F:enzyme regulator activity; IEA.

DR GO; GO:0006808; P:regulation of nitrogen utilization; IEA.

DR InterPro; IPR002187; PII_glnB.

DR Pfam; PF00543; P-II; 1.

DR PRINTS; PR00340; PIIGLNB.

DR ProDom; PD001194; PII_glnB; 1.

DR PROSITE; PS00638; PII_GlnB_CTER; 1.

SQ SEQUENCE 196 AA; 21275 MW; FE740EA66776F157 CRC64;

Query Match 91.4%; Score 509; DB 10; Length 196;

Best Local Similarity 90.3%; Pred. No. 8.9e-41;

Matches 102; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 FYKVEAILRPWRVSVSSALLKIGIRGVTVDVRGFGAQQGSTERQGGSEFSEDKFVAKV 60

Db 74 FYKVEAILRPWRVSVSSALLKIGIRGVTVDVRGFGAQQGSTERHGGSEFSEDKFVAKV 133

QY 61 KMEIVWSKDQVEDVIEKIIIEARTGEIGDGKIFLLPVSVDVIRVVTGERGDKAE 113

Db 134 KMEIVWSKDQVEDVIEKIIIEARTGEIGDGKIFLLPVSVDVIRVVTGERGDKAE 186

RESULT 3

Q9ARI4 PRELIMINARY; PRT; 194 AA.

AC Q9ARI4; 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE PII protein.

GN GLNB.

OS Medicago sativa (Alfalfa).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

OX NCBI_TaxID=3879;

RN [1]

RP SEQUENCE FROM N.A.

RA Garcia-Ibáñeta D., Sengupta-Gopalan C.;

RT "Characterization of PII (GLNB) in alfalfa.";

RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AY027892; AAK16221.1; -

DR HSSP; P38504; 1GNK.

DR GO; GO:0030234; F:enzyme regulator activity; IEA.

DR GO; GO:0006808; P:regulation of nitrogen utilization; IEA.

DR InterPro; IPR002187; PII_glnB.

DR Pfam; PF00543; P-II; 1.

DR PRINTS; PR00340; PIIGLNB.

DR ProDom; PD001194; PII_glnB; 1.

DR PROSITE; PS00638; PII_GlnB_CTER; 1.

SQ SEQUENCE 194 AA; 21413 MW; 23FA623FF4D97450 CRC64;

Query Match 88.3%; Score 492; DB 10; Length 194;

Best Local Similarity 85.8%; Pred. No. 3.7e-39;

Matches 97; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 1 FYKVEAILRPWRVSVSSALLKIGIRGVTVDVRGFGAQQGSTERQGGSEFSEDKFVAKV 60

Db 70 FYKVEAILRPWRVSVSSALLKIGIRGVTVDVRGFGAQQGSTERHGGSEFSEDKFVAKV 129

QY 61 KMEIVWSKDQVEDVIEKIIIEARTGEIGDGKIFLLPVSVDVIRVVTGERGDKAE 113

Db 130 KMEIVWSKDQVEDVIEKIIIEARTGEIGDGKIFLLPVSVDVIRVVTGERGDKAE 182

RESULT 4

Q9L422 PRELIMINARY; PRT; 112 AA.

AC Q9L422; 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE PII protein (Nitrogen regulatory protein P-II).

GN GLNB OR ALL2319.

OS Anabaena sp. (strain PCC 7120).

OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.

OX NCBI_TaxID=103690;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PCC 7120;

RA Gonzalez L., Phalip V., Zhang C.C.;

DR	EMBL; AE002548; AAF42322.1; -
DR	PIR; B81019; B81019.
DR	HSSP; P05826; 2PII.
DR	TIGR; NMB1995; -
DR	GO; GO:0030234; F:enzyme regulator activity; IEA.
DR	GO; GO:0006808; P:regulation of nitrogen utilization; IEA.
DR	InterPro; IPR002187; PII_glnB.
DR	InterPro; IPR002332; PII_glnB_UMP_S.
DR	Pfam; PF00543; P-II; 1.
DR	PRINTS; PR00340; PIIGLNB.
DR	ProDom; PD001194; PII_glnB; 1.
DR	PROSITE; PS00638; PII_GLNB_CTER; 1.
DR	PROSITE; PS00496; PII_GLNB_UMP; 1.
KW	Complete proteome.
SQ	SEQUENCE 112 AA; 12311 MW; 22963CF2526D4332 CRC64;
Query Match 54.3%; Score 302.5; DB 16; Length 112;	
Best Local Similarity 53.6%; Pred. No. 2.5e-21;	
Matches 59; Conservative 26; Mismatches 24; Indels 1; Gaps 1;	
QY	3 KVEAILRPWRVSVSSALLKIGIRGVTVDVRGFGAQQGSTERQGGSEFSEDKFAVKVM 62
Db	3 KIEAIVKPKLDDVREALTEIGITGMTVSEVKFGGRQKHTIYRGAEYAVD-FLPKIKI 61
QY	63 EIVVSKDQVEDVIEKIIIEARTGEIGDGKIFLLPVSVDVIRVTGERGOKA 112
Db	62 ELVLADDAVERAIDVIVEVARSGKIGDGKIFLPLVVEEAIRITGERSDAA 111
RESULT 9	
QY	QTV5R4 PRELIMINARY; PRT; 112 AA.
AC	QTV5R4;
DT	01-OCT-2003 (TrEMBLrel. 25, Created)
DT	01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	Nitrogen regulatory protein P-II.
GN	GLNB OR PMT1481.
OS	Prochlorococcus marinus (strain MIT 9313).
OC	Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC	Prochlorococcus.
OX	NCBI_TaxID=74547;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=22825698; PubMed=12917642;
RA	Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,
RA	Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA	Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
RA	Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA	Webb E.A., Zinser E.R., Chisholm S.W.;
RT	"Genome divergence in two Prochlorococcus ecotypes reflects oceanic
RT	niche differentiation."
RL	Nature 424:1042-1047(2003).
KW	EMBL; BX572099; CAE21656.1; -
KW	Complete proteome.
SQ	SEQUENCE 112 AA; 12362 MW; 5CA64D6663ED3B65 CRC64;
Query Match 54.3%; Score 302.5; DB 16; Length 112;	
Best Local Similarity 56.4%; Pred. No. 2.5e-21;	
Matches 62; Conservative 19; Mismatches 28; Indels 1; Gaps 1	
QY	3 KVEAILRPWRVSVSSALLKIGIRGVTVDVRGFGAQQGSTERQGGSEFSEDKFAVKVM 62
Db	3 KVEAIVRPFKLEDVKLALVNAEIIGMTVSEVRGFRQKGOVERVRCSEPTVE-FLQKLKV 61
QY	63 EIVVSKDQVEDVIEKIIIEARTGEIGDGKIFLLPVSVDVIRVTGERGOKA 112
Db	62 EVTVDDDKVEAVVNAIAEAAKTGEIGDGKIFISPDVSVVIRITGERDOKA 111
RESULT 10	
QY	Q8GQS4 PRELIMINARY; PRT; 112 AA.
IN	Q8GQS4

Fri May 28 14:57:10 2004

```

DR PRINTS; PR00340; PIIGLNB.
DR PRODOM; PD001194; PII_glnb; 1.
DR PROSITS; PS00638; PII_GLNB_CTER; 1.
DR PROSITE; PS00496; PII_GLNB_UMP; 1.
DR PROSITE; PS00496; PII_GLNB_UMP; 1.
FT NON TER 1
SO SEQUENCE 108 AA; 11812 MW; BE98D250C543ACBD CRC64;

```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 13:35:34 ; Search time 2414.89 Seconds
(without alignments)
14663.718 Million cell updates/sec

Title: US-09-756-541-13

Perfect score: 817

Sequence: 1 CTGAAAGTTGTGTTAAAAA.....GTTCTTAATAAAAAA 817

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	817	100.0	817	6	AR125589	AR125589 Sequence
2	808	98.9	844	8	AF095455	AF095455 Arabidops
3	591	72.3	591	6	AX507438	AX507438 Sequence
4	591	72.3	591	8	BT005209	BT005209 Arabidops
5	588	72.0	588	6	AR125591	AR125591 Sequence
6	286.4	35.1	897	6	AR125590	AR125590 Sequence
7	285	34.9	594	6	AR125592	AR125592 Sequence
8	285	34.9	840	8	AF095454	AF095454 Ricinus c
9	253.8	31.1	796	8	AY027892	AY027892 Medicago
10	246	30.1	120185	8	AC007138	AC007138 Arabidops
11	246	30.1	198220	8	ATCHRIV5	AL161493 Arabidops
12	245	30.0	947	8	AY442185	AY442185 Lycopersi
13	209.6	25.7	1085	8	PPI489604	AJ489604 Pinus pin
14	199.4	24.4	902	8	AK068407	AK068407 Oryza sat
15	199.4	24.4	902	8	AK099152	AK099152 Oryza sat
16	109.8	13.4	384	1	FDGLNBPRT	X97327 F.diplosiph
17	107.4	13.1	300143	1	AE017165	AE017165 Prochloro
18	106.8	13.1	164921	8	AF022186	AF022186 Cyanidium
19	106	13.0	2493	1	ASP251822	AJ251822 Anabaena
20	106	13.0	341880	1	AP003589	AP003589 Nostoc sp
21	102	12.5	262202	1	EX572094	EX572094 Prochloro
22	101.6	12.4	191028	8	PPU38804	U38804 Porphyra pu
23	98.6	12.1	4348	1	AF079137	AF079137 Synecococ
24	98	12.0	686	1	AF017419	AF017419 Nostoc pu
25	97	11.9	383	1	SYOGLNB	M62447 Synecococc
26	95.8	11.7	339	1	PMA271089	AJ271089 Prochloro
27	95.8	11.7	11771	1	AE000674	AE000674 Aquifex a
28	95.2	11.7	302320	1	AP005081	AP005081 Vibrio pa
29	94.2	11.5	349746	1	EX572099	EX572099 Prochloro
30	94	11.5	299350	1	AP005370	AP005370 Thermosyn
31	93.6	11.5	302050	1	AP006568	AP006568 Gloebact
32	93.2	11.4	300045	1	AE016803	AE016803 Vibrio vu
33	92.8	11.4	96109	6	AR408756	AR408756 Sequence
34	92.8	11.4	96109	6	AX067460	AX067460 Sequence
35	91.6	11.2	247950	1	AP005340	AP005340 Vibrio vu
36	89.8	11.0	11371	1	AE006236	AE006236 Pasteurel
37	89.4	10.9	351	6	AR378407	AR378407 Sequence
38	89.4	10.9	300732	1	AE016802	AE016802 Vibrio vu
39	89	10.9	1622	1	AF120107	AF120107 Synecococ
40	88.4	10.8	452	1	SSGLNBP	X97496 Synecocyst
41	88.4	10.8	130001	1	D90915	D90915 Synecocyst
42	86.8	10.6	298950	1	AP004597	AP004597 Oceanobac
43	86.2	10.6	248650	1	AP005341	AP005341 Vibrio vu
44	85.8	10.5	375	6	AR318761	AR318761 Sequence
45	85.4	10.5	405	6	AR376903	AR376903 Sequence

ALIGNMENTS

RESULT 1
AR125589
LOCUS AR125589 817 bp DNA
DEFINITION Sequence 13 from patent US 6177275.
ACCESSION AR125589
VERSION AR125589.1 GI:14111651
KEYWORDS
SOURCE Unknown.
ORGANISM Uncl. known.
REFERENCE 1 (bases 1 to 817)
AUTHORS Coruzzi, G.M., Lam, H.-M. and Hsieh, M.-H.
TITLE Plant nitrogen regulatory P-II genes
JOURNAL Patent: US 6177275-A 13 23-JAN-2001;
FEATURES Location/Qualifiers

Linear PAT 16-MAY-2001

[illegible]

RESULT 2
AF095455
LOCUS AF095455 844 bp mRNA linear PLN 19-NOV-1998
DEFINITION Arabidopsis thaliana PII protein mRNA, complete cds.
ACCESSION AF095455
VERSION AF095455.1 GI:3885942
KEYWORDS

```

SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 844)
Hsieh,M.H., Lam,H.M., van de Loo,F.J. and Coruzzi,G.
A PII-like protein in Arabidopsis: putative role in nitrogen
sensing
Proc. Natl. Acad. Sci. U.S.A. 95 (23), 13965-13970 (1998)
99030678
98111909
2 (bases 1 to 844)
Hsieh,M.H., Lam,H.M., van de Loo,F.J. and Coruzzi,G.M.
Direct Submission
Submitted (29-SEP-1998) Biology, New York University, 100
Washington Square East, New York, NY 10003, USA
Location/Qualifiers
1. .844
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="IV"
/map="10.8 CM"
31. .621
/function="putative role in nitrogen sensing"
/note="GLP1"
/codon_start=1
/product="PII protein"
/protein_id="AAC78333.1"
/db_xref="GI:3885943"
/translation="MAASMTKPISITSLGFSYDRKNIAFSDCIS:CSGFRHSRPSCLD
LVTQSNNSRVLPPWSAQISSDYIPDSKPYKVEIVRPRIQQVSSALLKIGIRGV
VSDVRGFGAQQGSTERHGGSEFSEDKFVAKVKMEIVWKQDVESVINTIIEGARTGEI
GDGKIFVLTPSDIVIRVTGERGEKAEMKMGDMLSPS"

CDS

```


QY	483	ATAATTGAAGGAGCAAGGACAGAGAGATTGGTGTATGGCAAGATTTTGTGTTTGCCTGTG	542
Db	481	ATAATTGAAGGAGCAAGGACAGAGAGATTGGTGTATGGCAAGATTTTGTGTTTGCCTGTG	540
QY	543	TCAGATGTCATAAGAGTTAGGACAGGTGAGCGTGGGGAGAAAGCAGAGAGATGACTGGT	602
Db	541	TCAGATGTCATAAGAGTTAGGACAGGTGAGCGTGGGGAGAAAGCAGAGAGATGACTGGT	600
QY	603	GATATGCTTACCGTCTTAGGAACAACAGAGCTCAAGAAATGGTTTATTTTTCAT	662
Db	601	GATATGCTTACCGTCTTAGGAACAACAGAGCTCAAGAAATGGTTTATTTTTCAT	660
QY	663	TTGGGTCTTAGATCTCGGAATAATAATGAATGAGTCTGTGTTTGGTTTCATGTTGAA	722
Db	661	TTGGGTCTTAGATCTCGGAATAATAATGAATGAGTCTGTGTTTGGTTTCATGTTGAA	720
QY	723	TCGATCAAGATGTGTTTTTAACCTGTACATGAATATGACAGAAACATCTCTCTGTTCTC	782
Db	721	TCGATCAAGATGTGTTTTTAACCTGTACATGAATATGACAGAAACATCTCTCTGTTCTC	780
QY	783	AGACATCGAAACTCTGTTTCCTAATAAAA	810
Db	781	AGACATCGAAACTCTGTTTCCTAATAAAA	808
RESULT 3	AX507438	LOCUS	AX507438
DEFINITION	Sequence 2133 from Patent WO0216655.	591 bp	DNA
ACCESSION	AX507438	linear	PAT 27-SEP-2002
VERSION	AX507438.1	GI:23388675	
KEYWORDS	Arabidopsis thaliana (thale cress)		
SOURCE	Arabidopsis thaliana		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1	Harper, J.F., Kreps, J., Wang, X. and Zhu, T.	
AUTHORS		Stress-regulated genes of plants, transgenic plants containing	
TITLE		same, and methods of use	
JOURNAL		Patent: WO 0216655-A 2133 28-FEB-2002;	
		The Scripps Research Institute (US); Syngenta Participations AG	
		(CH)	
FEATURES	Location/Qualifiers		
source	1..591		
	/organism="Arabidopsis thaliana"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:3702"		
ORIGIN			
	Query Match	72.3%;	Score 591; DB 6; Length 591;
	Best Local Similarity	100.0%;	Pred. No. 4.1e-130;
	Matches	591; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	33	ATGGCGGCGTCAATGACGAAACCCATCTCAATAACTTCTCGGTTTCTATTCTGATCGA	92
Db	1	ATGGCGGCGTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGA	60
QY	93	AAGAACATTGCTTTCTCTGATTGCAATTCGATTGTTCTGGAATTCAGACATTCGCGACCA	152
Db	61	AAGAACATTGCTTTCTCTGATTGCAATTCGATTGTTCTGGAATTCAGACATTCGCGACCA	120
QY	153	TCTTGGCTCGATTGGTCAACAAGTCACCGAGTAATAACAGTCGTCGTTTACCTGTCGTT	212
Db	121	TCTTGGCTCGATTGGTCAACAAGTCACCGAGTAATAACAGTCGTCGTTTACCTGTCGTT	180
QY	213	AGTGCCCAATATCTTCTGATTATATTCAGACTCGAAATTTTACAGGTGGAAGCAATT	272
Db	181	AGTGCCCAATATCTTCTGATTATATTCAGACTCGAAATTTTACAGGTGGAAGCAATT	240
QY	273	GTCAGACCATGGAGAAATCCAGCAAGTTTCATCGGCTTACTGAAATTCGGGATTCGAGGT	332

Db	241	GTCAGACCATGGAGAAATCCAGCAAGTTTTCATCGGCTTTTACTGAAATCCGGATTCGAGGT	300
QY	333	GTTACTGTTTCTGATGTGAGAGGGTTTGGTGTGCACAAGGAGGTTTCTACCGAGAGACACGGT	392
Db	301	GTTACTGTTTCTGATGTGAGAGGGTTTGGTGTGCACAAGGAGGTTTCTACCGAGAGACACGGT	360
QY	393	GGCTCTGAGTTCTCGGAAGACAAATTTGTTGTCTAAAGTTAAGATGGAATCGTTGTTAAG	452
Db	361	GGCTCTGAGTTCTCGGAAGACAAATTTGTTGTCTAAAGTTAAGATGGAATCGTTGTTAAG	420
QY	453	AAAGACCAAGTGGAAATCTGTAATCAACACACATAATTTGAAGGAGCAAGGACAGAGAGATT	512
Db	421	AAAGACCAAGTGGAAATCTGTAATCAACACACATAATTTGAAGGAGCAAGGACAGAGAGATT	480
QY	513	GGTGATGGCAAGATTTTGTGTTTGCCTGTGTACATGTATAGAGTTAGGACAGGTGAG	572
Db	481	GGTGATGGCAAGATTTTGTGTTTGCCTGTGTACATGTATAGAGTTAGGACAGGTGAG	540
QY	573	CGTGGGAGAAAGCAGAGAGATGACTGGTGATATGCTTTTACCCTCTTAG	623
Db	541	CGTGGGAGAAAGCAGAGAGATGACTGGTGATATGCTTTTACCCTCTTAG	591

RESULT 4

BT005209	LOCUS	BT005209	591 bp	mRNA	linear	PLN 14-MAR-2003
DEFINITION	Arabidopsis thaliana At4g01900 mRNA, complete cds.					
ACCESSION	BT005209					
VERSION	BT005209.1	GI:28950698				
KEYWORDS	FLI CDNA.					
SOURCE	Arabidopsis thaliana (thale cress)					
ORGANISM	Arabidopsis thaliana					
REFERENCE	1	(bases 1 to 591)				
AUTHORS	Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Hsuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.					
TITLE	Arabidopsis ORF clones					
JOURNAL	Unpublished					
REFERENCE	2	(bases 1 to 591)				
AUTHORS	Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Hsuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.					

TITLE

Direct Submission
Submitted (14-MAR-2003) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Chan, M.M., Chang, C.M., Dale, J.M., Hsuan, V.W., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and

Ecker, J.R.

Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

source

1. .591

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="4"

/clone="U23463"

/ecotype="Columbia"

/note="This clone is in PUNI 51"

1. .591

/note="p II nitrogen sensing protein GLB I"

/codon_start=1

/product="At4g01900"

/protein_id="AA063273.1"

/db_xref="GI:28950699"

/translation="MAASMTKPIISITSLGFYSDRKNIAFSDCISICSGFRHSRPSCLD LVTKSPNSRVLPPVSAQISSDYIPDSKFYKVAIVRPWRIQQVSSALLKIGIRGVT VSDVRGFGAQGGSTHRHGGSEFSEDKFVAKVMEIVKKQDQVESVINTIIEGARTGEI GDGKIFVLPSVDIVRVTGERGEKAEKMTGDMLSPS"

ORIGIN

Query Match 72.3%; Score 591; DB 8; Length 591;

Best Local Similarity 100.0%; Pred. No. 4.1e-130; Mismatches 0; Indels 0; Gaps 0;

Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 ATGGCGCGTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGA 92

Db 1 ATGGCGCGTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGA 60

QY 93 AAGAACATTGCTTTCTCTGATTGCATTTCTGATTGTTCTGATTTCAGACATTCCTCGACCA 152

Db 61 AAGAACATTGCTTTCTCTGATTGCATTTCTGATTGTTCTGATTTCAGACATTCCTCGACCA 120

QY 153 TCTTGCCCTCGATTGGTTCACAAAGTCACGAGTAAATACAGTCTGTTTACCTGCTT 212

Db 121 TCTTGCCCTCGATTGGTTCACAAAGTCACGAGTAAATACAGTCTGTTTACCTGCTT 180

QY 213 AGTGCCCAATATCTTCTGATTATATTCAGACTCGAAATTTTACAGGTGGAAGCAATT 272

Db 181 AGTGCCCAATATCTTCTGATTATATTCAGACTCGAAATTTTACAGGTGGAAGCAATT 240

QY 273 GTCCAGCATGGAGAAATCCAGCAAGTTTCTATCGGCTTTACTGAAATCGGGATTCGAGGT 332

Db 241 GTCCAGCATGGAGAAATCCAGCAAGTTTCTATCGGCTTTACTGAAATCGGGATTCGAGGT 300

QY 333 GTTACTGTTTCTGATGTGAGAGGGTTTGGTGCAAGGAGGTTCTACCGAGACACCGGT 392

Db 301 GTTACTGTTTCTGATGTGAGAGGGTTTGGTGCAAGGAGGTTCTACCGAGACACCGGT 360

QY 393 GGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTTAAAGTTAAGATGGAATCGTTTAAAG 452

Db 361 GGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTTAAAGTTAAGATGGAATCGTTTAAAG 420

QY 453 AAAGACCAAGTGGAAATCTGTAATCAACATAATTGAAGGACCAAGGACGAGAGATT 512

Db 421 AAAGACCAAGTGGAAATCTGTAATCAACATAATTGAAGGACCAAGGACGAGAGATT 480

QY 513 GGTGATGGCAAGATTTTGTGTTTGCCTGTGTCAGATGTCATAAGATTAGGACAGGTGAG 572

Db 481 GGTGATGGCAAGATTTTGTGTTTGCCTGTGTCAGATGTCATAAGATTAGGACAGGTGAG 540

QY 573 CGTGGGAGAAAGCAGAGAGATGACTGGTGATATGCTTTCACCGTCTTAG 623

Db 541 CGTGGGAGAAAGCAGAGAGATGACTGGTGATATGCTTTCACCGTCTTAG 591

RESULT 5

AR125591

LOCUS

Sequence 15 from patent US 6177275.

Accession AR125591

Version AR125591.1

Keywords

Source

Unknown.

Organism

Unclassified.

REFERENCE

1 (bases 1 to 588)

Coruzzi, G.M., Lam, H.-M. and Hsieh, M.-H.

Plant nitrogen regulatory P-PII genes

Patent: US 6177275-A 15 23-JAN-2001;

Location/Qualifiers

1. .588

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match 72.0%; Score 588; DB 6; Length 588;

Best Local Similarity 100.0%; Pred. No. 2.1e-129; Mismatches 0; Indels 0; Gaps 0;

Matches 588; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 ATGGCGCGTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGA 92

Db 1 ATGGCGCGTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGA 60

QY 93 AAGAACATTGCTTTCTCTGATTGCATTTCTGATTGTTCTGATTTCAGACATTCCTCGACCA 152

Db 61 AAGAACATTGCTTTCTCTGATTGCATTTCTGATTGTTCTGATTTCAGACATTCCTCGACCA 120

QY 153 TCTTGCCCTCGATTGGTTCACAAAGTCACGAGTAAATACAGTCTGTTTACCTGCTT 212

Db 121 TCTTGCCCTCGATTGGTTCACAAAGTCACGAGTAAATACAGTCTGTTTACCTGCTT 180

QY 213 AGTGCCCAATATCTTCTGATTATATTCAGACTCGAAATTTTACAGGTGGAAGCAATT 272

Db 181 AGTGCCCAATATCTTCTGATTATATTCAGACTCGAAATTTTACAGGTGGAAGCAATT 240

QY 273 GTCCAGCATGGAGAAATCCAGCAAGTTTCTATCGGCTTTACTGAAATCGGGATTCGAGGT 332

Db 241 GTCCAGCATGGAGAAATCCAGCAAGTTTCTATCGGCTTTACTGAAATCGGGATTCGAGGT 300

QY 333 GTTACTGTTTCTGATGTGAGAGGGTTTGGTGCAAGGAGGTTCTACCGAGACACCGGT 392

Db 301 GTTACTGTTTCTGATGTGAGAGGGTTTGGTGCAAGGAGGTTCTACCGAGACACCGGT 360

QY 393 GGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTTAAAGTTAAGATGGAATCGTTTAAAG 452

Db 361 GGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTTAAAGTTAAGATGGAATCGTTTAAAG 420

QY 453 AAAGACCAAGTGGAAATCTGTAATCAACATAATTGAAGGACCAAGGACGAGAGATT 512

Db 421 AAAGACCAAGTGGAAATCTGTAATCAACATAATTGAAGGACCAAGGACGAGAGATT 480

QY 513 GGTGATGGCAAGATTTTGTGTTTGCCTGTGTCAGATGTCATAAGATTAGGACAGGTGAG 572

Db 481 GGTGATGGCAAGATTTTGTGTTTGCCTGTGTCAGATGTCATAAGATTAGGACAGGTGAG 540

QY 573 CGTGGGAGAAAGCAGAGAGATGACTGGTGATATGCTTTCACCGTCTTAG 620

Db 541 CGTGGGAGAAAGCAGAGAGATGACTGGTGATATGCTTTCACCGTCTTAG 588

RESULT 6

AR125590

LOCUS

Sequence 14 from patent US 6177275.

Accession AR125590

Version AR125590.1

Keywords

Source

Unknown.

Organism

Unclassified.

REFERENCE 1 (bases 1 to 897)
AUTHORS Coruzzi,G.M., Lam,H.-M. and Hsieh,M.-H.
TITLE Plant nitrogen regulatory P-PII Genes
JOURNAL Patent: US 617275-A 14 23-JAN-2001;
FEATURES Location/Qualifiers
source 1..897
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 35.1%; Score 286.4; DB 6; Length 897;
Best Local Similarity 69.8%; Pred. No. 1.3e-57;
Matches 402; Conservative 0; Mismatches 171; Indels 3; Gaps 1;

QY 28 GAATCATGGCGGCTCAATGACGAAACCCATCTCAATAAAGTCTCTCGGTTTCTATTTCTG 87
DB 45 GAATTCGGCAGGAGCTACTGCGAAACTGGGCTTGCTCACTCTCTTCAATCTAATAACA 104
QY 88 ATCGAAAGAACATTGCTTCTCTGATTGCAATTCGATTTGTTCTGGATTTCAGACATCCC 147
DB 105 TCAAGAAAGAAATTCCTGTTTTGATTTTCAGTTTGTTGTTGCCAGAGCTTAGACATTCTC 164
QY 148 GACCATCTTGCTCGAATTGGTCACAAAGTCAACCGAGTAATAACAGTCTGTTTACCTG 207
DB 165 GGTTTTCTCACTTTAAACACCGCGGTCAAGCGGTAAAGATATGCC---CCGTCGTTCTG 221
QY 208 TCGTTAGTGCCCAATATCTTCTGATTATATTCAGACTCGAAATTTTACAAGGTGGAAG 267
DB 222 TGATTAAATGCCCAAGCTCGCCTGACTACATTCCTGATGCTAAATTTCTACAAAGTGAAG 281
QY 268 CAATTGTCAGACCATGAGAAATCCAGCAAGTTTCATCGGCTTTACTGAAATTCGGGATTC 327
DB 282 CAATTCTCAGGCCCTGGCGAGTCTCGCAAGTTTCTCGGCTTTGCTAATAAATTGGTATTC 341
QY 328 GAGGTGTTACTGTTTCTGATGTGAGAGGTTTGGTGCACAGGAGGTTCTACCGAGAGAC 387
DB 342 GAGGTGTTACTGTTTCTGATGTTTCGAGGTTTGGTGTCTCAAGGTGGTTCAACTGAGAGGC 401
QY 388 ACGGTGGCTCTGAGTTCCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATCGAATCGTTG 447
DB 402 AGGCGGCTCAGAAATTTTCTGAAGACAAAGTTTGTGCTAAAGTTAAGATCGGATCGTGG 461
QY 448 TTAAGAAAGACCAAGTGGAATCTGTAATCAACACAATAAATGAAGCAAGGACAGGAG 507
DB 462 TTAGCAAGAGACCAAGTTTGAGGATGTTATAGAAAATAATCAATTGAGGAGGCAAGAACTGGAG 521
QY 508 AGATTGGTGATGGCAAGATTTTGTGTTTTCCTGTGTGTCAGATGTCATAAGAGATTAGGACAG 567
DB 522 AGATTGGAGACGGCAAGATTTTCTGCTGCTGTTTTCAGATGTAATAAGAGTCCGCACCTG 581
QY 568 GTGAGCGTGGGAGAAAGCAGAGAGAGATGACTGGTG 603
DB 582 GTGAGCGGGTGATAGAGCTGAGAGGATGACAGGAG 617

RESULT 7
AR125592
LOCUS AR125592 594 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 16 from patent US 617275.
ACCESSION AR125592
VERSION AR125592.1 GI:14111654
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 594)
TITLE Coruzzi,G.M., Lam,H.-M. and Hsieh,M.-H.
JOURNAL Plant nitrogen regulatory P-PII Genes
FEATURES Patent: US 617275-A 16 23-JAN-2001;
source Location/Qualifiers
1..594
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 34.9%; Score 285; DB 6; Length 594;
Best Local Similarity 70.3%; Pred. No. 2.8e-57;
Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

QY 39 GCGTCAATGACGAAACCCATCTCAATAAAGTCTCTCGGTTTCTATTCTGATCGAAAGAAC 98
DB 7 GAGGCTACTGCGAAACTGGGCTTGCTCACTCTCTTCAATCTAATAACATCAAGAAAGAA 66
QY 99 ATTGCTTTTCTCTGATTGCAATTCGATTTGTTCTCGATTTCAGACATTCCTCCGACCATCTTGC 158
DB 67 TTCCCTGTTTGGATTTCAGTTTGTTTGTCCAGAGCTTAGACATTTCTCGGTTTCTCAC 126
QY 159 CTCGATTGGTCAAAAGTCAACGAGTACCGAGTAATAACAGTCTGTTTACCTGTCTGTAGTGC 218
DB 127 TTTAACACCGCGTCAAGCGGTAAGATATGCC---CCCGTCGTTCTGTGATTAAATGCC 183
QY 219 CAAATATCTTCTGATTATATTCAGACTCGAAATTTTACAGGTGGAAGCAATTTGTCAGA 278
DB 184 CAAAGCTCGCCTGACTACATTCCTGATGCTAAATTTCTACAAAGTGAAGCAATTTCTCAGG 243
QY 279 CCATGGAGAATCCAGCAAGTTTCATCGGCTTTACTGAAAATCGGATTCGAGGTGTTACT 338
DB 244 CCCTGGCGAGTCTCGCAAGTTTCTCTCGGCTTTGCTAATAAATTTGGTATTCGAGGTGTTACT 303
QY 339 GTTCTGATGTGAGAGGTTTGGTGCAAGGAGGTTCTACCGAGAGACACGGTGGCTCT 398
DB 304 GTTCTGATGTTCGAGGTTTGGTGTCTCAAGGTGGTTCAACTGAGAGGCGAGGCGGCTCA 363
QY 399 GAGTTCTCGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAGAAAGAC 458
DB 364 GAATTTCTGAAGACAAAGTTTGTCTAAAGTTAAGATGGAGATCGTGGTTAGCAAGAC 423
QY 459 CAAGTGGAAATCTGTAATCAACACAATAATTAAGAGGACGAGGAGAGATTTGGTGAT 518
DB 424 CAGGTTGAGGATGTTATAGAAAATAATCAATTGAGGAGGCAAGAACTGGAGAGATTGGAGAC 483
QY 519 GGCAGATTTTGTGTTTGCCTGTGTCAAGTGTCTAAGAGTTAGGACAGGTGAGCGTGG 578
DB 484 GGCAGATTTTCTTGCTGCCGTTTTCAGATGTAATAAGAGTCCGCACCTGGTGAGCGGGT 543
QY 579 GAGAAAGCAGAGAGAGATGACTGGTG 603
DB 544 GATAAGGCTGAGAGGATGACAGGAG 568

RESULT 8
AF095454
LOCUS Ricinus communis PII protein mRNA, partial cds.
DEFINITION Ricinus communis PII protein mRNA, partial cds.
ACCESSION AF095454
VERSION AF095454.1 GI:3885940
KEYWORDS
SOURCE Ricinus communis (castor bean)
ORGANISM Ricinus communis
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Malpighiales; Euphorbiaceae; Acalyphoideae; Acalypheae; Ricinus.
AUTHORS 1 (bases 1 to 840)
TITLE Hsieh,M.H., Lam,H.M., van de Loo,F.J. and Coruzzi,G.
JOURNAL A PII-like protein in Arabidopsis: putative role in nitrogen sensing
MEDLINE Proc. Natl. Acad. Sci. U.S.A. 95 (23), 13965-13970 (1998)
PUBMED 99030678
REFERENCE 2 (bases 1 to 840)
AUTHORS Hsieh,M.H., Lam,H.M., van de Loo,F.J. and Coruzzi,G.M.
TITLE Direct Submission
JOURNAL Submitted (29-SEP-1998) Biology, New York University, 100 Washington Square East, New York, NY 10003, USA
FEATURES Location/Qualifiers

source	1. .840	/organism="Ricinus communis"	
		/mol_type="mRNA"	
		/db_xref="taxon:3988"	
CDS	<1. .595	/function="putative role in nitrogen sensing"	
		/note="GLB1"	
		/codon_start=2	
		/product="PII protein"	
		/protein_id="AAC78332.1"	
		/db_xref="GI:3885941"	
		/translation="HEATAKLGILLTPHNSNNIKKEFPVPDFSLFCPELRHRSRFSHFNT	
		AVKRVYAPVPVINAQSPDYIPDAKFYKVEAILRPWRVSVSSALLKIGIRGVTVS	
		DVRGFGAQQGSTERQGGSEFSEDKFVAKVMEIVVSKDQVEDVIEKIIIEARTGEIGD	
		GKIFLLPVSDVIRVTSERGDKAERMTGGRSDMSTA"	
ORIGIN			
	Query Match	34.9%;	Score 285; DB 8; Length 840;
	Best Local Similarity	70.3%;	Pred. No. 2.7e-57;
	Matches	397; Conservative	0; Mismatches 165; Indels 3; Gaps 1;
QY	39	GGGTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGAAGAAC	98
Db	5	GAGGCTACTGCGAAACTGGCTTGCTCACTCCTCTTCAATCTAATAACATCAAGAAGAA	64
QY	99	ATTGCTTCTCTGATTGCATTTCGATTGTTCTGGATTTCAGACATTCGCGACCATCTTGC	158
Db	65	TTCCCTGTTTGAATTCAGTTTGTGTTTGTGCCAGACTTAGACATTCCTCGTTTCTCAC	124
QY	159	CTCGATTGTTGTCACAAAGTCACCGAGTAATAACAGTCTGTTTTTACCTGTCGTTAGTGCC	218
Db	125	TTTAACACCGCGGTCAAGCGGTAAGATATGCC---CCCGTCGTTCTTGATTAATGCC	181
QY	219	CAAAATATCTCTGATTATATTCAGACTCGAAATTTTACAAGGTGGAAGCAATGTGCAGA	278
Db	182	CAAGCTCGCCTGACTACATTCCTGATGCTAAATTTCTACAAGTGAAGCAATTCCTCAGG	241
QY	279	CCATGGAGATCCAGCAAGTTTCATCGGCTTTACTGAAATTCGGGATTCGAGGTGTTACT	338
Db	242	CCCTGGGAGTCTCGCAAGTTTCTCGGCTTTGCTAAGAAATTTGGTATTCGAGGTGTTACT	301
QY	339	GTTTCTGATGTGAGAGGGTTTGGTGACAAAGGAGGTTTCTACCGAGAGACACGGTGGCTCT	398
Db	302	GTTTCTGATGTTCGAGGTTTGGTGCTCAAGGTGTTCAACTGAGAGGCGAGGCGGCTCA	361
QY	399	GAGTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAGAAAGAC	458
Db	362	GAATTTCTGAAGACAAGTTTGTGCTAAAGTTAAGATGGAGATCGTGGTTAGCAAAGAC	421
QY	459	CAAGTGGAACTCTGTAATCAACACAAATATTGAAGGAGCAAGGACAGAGAGATTGGTGAT	518
Db	422	CAGGTTGAGGATGTTATAGAAAAAATCATTTGAGGAGGCAAGAACTGGAGAGATTGGAGAC	481
QY	519	GGCAAGATTTTGTGTTGCTGCTGTCAGATGTCATAAGATTAGGACAGGTGAGCGTGGG	578
Db	482	GGCAAGATTTTCTTGTGCTGCTGTTTCAGATGTAATAAGATCCGCACTGGTGAGCGGGT	541
QY	579	GAGAAAGCAGAGAAGATGACTGGTG	603
Db	542	GATAAGGCTGAGAGGATGACAGGAG	566
RESULT 9			
LOCUS	AY027892	796 bp	mRNA linear PLN 12-MAR-2001
DEFINITION	Medicago sativa PII protein (GLNB) mRNA, complete cds.		
ACCESSION	AY027892		
VERSION	AY027892.1	GI:13277514	
KEYWORDS			
SOURCE	Medicago sativa		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		

rosids; eurosids 1; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.			
1 (bases 1 to 796)			
Garcia-Ibailcieta, D. and Sengupta-Gopalan, C.			
Characterization of PII (GLNB) in alfalfa			
Unpublished			
2 (bases 1 to 796)			
Garcia-Ibailcieta, D. and Sengupta-Gopalan, C.			
Direct Submission			
Submitted (22-FEB-2001) Molecular Biology Program, New Mexico State University, Corner of Knox and College, Las Cruces, NM 88003, USA			
Location/Qualifiers			
1. .796			
/organism="Medicago sativa"			
/mol_type="mRNA"			
/db_xref="taxon:3879"			
1. .796			
/gene="GLNB"			
1. .585			
/gene="GLNB"			
/codon_start=1			
/product="PII protein"			
/protein_id="AAK16221.1"			
/db_xref="GI:13277515"			
/translation="MTLIAKPNVFNGLNFNINETQIPSSPSVIRKFRFGDSSHRNVVL			
RSNGNASVLPKIRAQNLDPVVPESKFYKVEAILRPWRIPIQVSSGLLKMGRGVTVSDV			
KGFQAQGSKERQGGSEFSEDNFVAKVMEIVVRKQDQVEAVINKIMETARTGEIGDSK			
IFLIPVSDVIRITGERSEQAERMAGGLTDALYV"			
ORIGIN			
Query Match	31.1%;	Score 253.8; DB 8; Length 796;	
Best Local Similarity	71.0%;	Pred. No. 7.3e-50;	
Matches	336; Conservative	0; Mismatches 137; Indels 0; Gaps 0;	
QY	198	GTTTACCTGCTTAGTGCCCAATATCTTCTGATTATATTCAGACTCGAAATTTTAC	257
Db	154	GTTCTTCCCAAAATTAGAGCTCAAAACCTTCTGACTATGTTCTCGAATCCAAATTTTAC	213
QY	258	AAGGTGGAAGCAATTGTCCAGACCATGGAGATCCAGCAAGTTTTCATCGGCTTTACTGAAA	317
Db	214	AAAGTTGAAGCCATTCTCAGGCCATGGAGATCCCTCAGGTTTCTTCGGGTTGTTGAAA	273
QY	318	ATCGGGATTCGAGGTGTTACTGTTTCTGATGTGAGAGGGTTTGGTGACAGAGGTTCT	377
Db	274	ATGGGAATTCGTGGTGTCACTGTATCTGATGTCAAGGGTTTGGTCTCAGGGTGGCTCA	333
QY	378	ACCGAGAGACACGGTGGCTCTGAGTTCTCGAAGACAAATTTGTTGCTAAAGTTAAGATG	437
Db	334	AAAGAGAGGCGAGGAGGCTCTGAATTTTCTGAAGACAATTTTGTGCCAAAGTTAAATG	393
QY	438	GAAATCGTTGTTAAGAAAGACCAAGTGGATCTGTAATCAACACAATAATTGAAGAGCA	497
Db	394	GAAATAGTGGTGAGAAAGACCAAGTTGAGGAGTGGAGAGTATAAAACAAATTTATGGAGCGCA	453
QY	498	AGGACAGGAGAGATTGGTGATGGCAAGATTTTGTGTTGCTGCTGTGTGATGATGATGATG	557
Db	454	AGAACTGGGAGATTGGTGATGGCAAAATTTTCTTGATCCCTGTATCTGATGTAATAAGA	513
QY	558	GTTAGGACAGGTGAGCTGGGGAGAAAGCAGAGAGATGACTGGTGATATGCTTTTCAAG	617
Db	514	ATCCGCACAGGTGAGCGTGGGGAGCAGGCTGAGAGATGGCTGGGGGACTAAGTGAAGCGG	573
QY	618	TCTTAGGAACAACAGAGCTCAAGAAATGGTTTTTTTTTTTTTTTTCATTTTGGTCT	670
Db	574	TTGTATGTTTGAACGAAAGATTAGCATTTTGCAATTTGGTCTCTGTTCTCTTT	626
RESULT 10			
LOCUS	AC007138	120185 bp	DNA linear PLN 01-APR-1999
DEFINITION	Arabidopsis thaliana BAC T7B11 from chromosome IV near 10 cM, complete sequence.		
ACCESSION	AC007138		

VERSION AC007138.1 GI:4510323

KEYWORDS HTC.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 120185)

AUTHORS Huang,E.N., Nascimento,L., de la Bastide,M., Habermann,K., Vil,M.D., Preston,R.R., Spiegel,L.A., See,L.H., Shah,R., Matero,A., O'Shaughnessy,A., Rodriguez,M., Shekher,M., Swaby,I., Schutz,K., Parnell,L.D., Dedhia,N.N. and McCombie,W.R.

TITLE Arabidopsis thaliana BAC T7B11 from chromosome IV near 10 cM

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 120185)

AUTHORS Huang,E.N., Nascimento,L., de la Bastide,M., Habermann,K., Vil,M.D., Preston,R.R., Spiegel,L.A., See,L.H., Shah,R., Matero,A., O'Shaughnessy,A., Rodriguez,M., Shekher,M., Swaby,I., Schutz,K., Parnell,L.D., Dedhia,N.N. and McCombie,W.R.

TITLE Direct Submission

JOURNAL Submitted (23-MAR-1999) Lita Annenberg Hazen Genome Sequencing Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724

REFERENCE 3 (bases 1 to 120185)

AUTHORS Parnell,L.D.

TITLE Direct Submission

JOURNAL Submitted (01-APR-1999) Lita Annenberg Hazen Genome Sequencing Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724

REMARK Arabidopsis thaliana BAC T7B11 from chromosome IV near 9.5 cM

COMMENT The T7 end of T7B11 is oriented toward the NOR and overlaps with T15B16, GenBank accession number AF104919. The SP6 end of T7B11 is oriented toward the centromere and overlaps with T10M13, GenBank accession number AF001308. There are no discrepancies found in these independently finished overlaps. The region from position 110064 to 110253 exists as single stranded and single chemistry. Please see the feature below for details.

FEATURES

source Location/Qualifiers

1..120185

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/cultivar="Columbia"

/db_xref="taxon:3702"

/chromosome="IV"

/map="near 10 cM"

misc_feature 1..14923

/note="overlap with T15B16, GenBank accession number AF104919, from position 1 to 14923"

gene 50..1880

/gene="T7B11.1"

/note="encodes hypothetical protein; identical to T15B16.9, GenBank accession number AF104919; gene model last edited on 3 Mar 99"

/evidence=not_experimental

join(50..556,1098..1259,1340..1561,1668..1880)

/gene="T7B11.1"

/note="identical to T15B16.9, GenBank accession number AF104919; similar to T7B11.3; functional catalog ID=99"

/codon_start=1

/evidence=not_experimental

/product="hypothetical protein"

/protein_id="AAD22637.1"

/db_xref="GI:4558544"

/translation="MAQKQOQLHQRPFSPPSSPRSYSSISNRPIFLSRNGLLLVLLAL FLLGLVLPWGPSPLLLFPNKVSSPSYASSLSHPAKSEWRNYTLAQAKFVATNGTVI VCAVSSPELPFLNNWLISVSRQKHQEKVLVIAEDYITLYKNEKMPGHAVLIPALDS KTAYSFGSGGFNTARRPQHLLQILELGVNVMYNDVDMVLQDPFOYLEGSHDAYFT DMPQIKPLNHSDDLPAPODNGETIYCSCMIYLRPTNGAKLLMKWSEELQSQAWSES IRFQANDQAPNALNKTAHQVDLYLLSQVAFPTGGLYFNDAAWVKETKGKHIIVHNN YIIGYDRKMRFPDYGGLWLVDHHALESPLGKLE"

complement(2647..4650)

/gene="T7B11.2"

/note="encodes putative CHP-rich zinc finger protein;

identical to T15B16.6, GenBank accession number AF104919; gene model last edited on 3 Mar 99"

/evidence=not_experimental

complement(2647..4650)

/gene="T7B11.2"

/note="identical to T15B16.6, GenBank accession number AF104919; similar to A. thaliana CHP-rich proteins encoded by T10M13, GenBank accession number AF001308; functional catalog ID=98"

/codon_start=1

/evidence=not_experimental

/product="putative CHP-rich zinc finger protein"

/protein_id="AAD22638.1"

/db_xref="GI:4558545"

/translation="MDSGVSVPILIEHLMPCNDLRRGCCERLEAISDGYCTTCD FFVHKCGDEASECIENPHSNHPLRLGFLSRLOQHWLKVRSCLDCKNGIDLFYR CEICDFMDLHCAKYPPEVIDIDPEMHSKLNLLKDRVEFDCDAKCGKIGYGFYECH ECDSKPHVDCVRYSSSEEVKHPLEVNHSYHSLHPLKLLMGQPPDYSDGKCRLCGRKID DKLPYHCSSCNFTLDMRCVLPNPPKSVLDLKTHTDHLNLLPRLIFFTCNACGLNGDRS PYACFCQDFLIHKDCFLRLINRHDHRISTRSLIGVNSVCGVCRQKVDWTCGGY SCORCSTYIAHSKCATREDVWNGKELEGWPEIEIDIEPVVDDNTIQHFSKHEHYMR LNVNGLMCEVNRKCNACNHPISQSPFYGCMDCDFILHQCAGFPRRKWHVILHNERLAL VTSVNIIFGCSACHIFNGFRVZHEDTKDLVLCGSFSEPIHPSHHPHPLYCISPEDD EVCSGCSNERSYHVLRCIEDNCGFILDGFCATFPQVVKHRIEDQFLSLCYGEKASGYW CDICEKTEPNTWFTCKDHRASLHTWCVLGDFMGLPKSTIELMNIYSYEVVLNNSIS RPICRHCKSHCIPPIILKKIGTSDPYFCSLDCIESEKRLWRAK"

8080..10112

/gene="T7B11.3"

/note="encodes hypothetical protein; identical to T15B16.8, GenBank accession number AF104919; gene model last edited on 3 Mar 99"

/evidence=not_experimental

join(8080..8568,9318..9479,9557..9778,9900..10112)

/gene="T7B11.3"

/note="identical to T15B16.8, GenBank accession number AF104919; similar to T7B11.1; functional catalog ID=99"

/codon_start=1

/evidence=not_experimental

/product="hypothetical protein"

/protein_id="AAD22639.1"

/db_xref="GI:4558546"

/translation="MEQKHILKQSTFSSSPSSYSSISDRPISILSRNGLLLLLALV LLLGLVLPWGPSPLFLFPNRLSSLSRSPQSKWRDYLTAQAARFVAKNGTVIVCAVSS PFLPFLNNWLISVSRQKHQKQVLIADYITLYKNEKWPGHAVLIPALDSKTAFSP GSQGFNTARRPQHLLQILELGVNVMYNDVDMVLQDPFLYLEGSHDAYFTDDMPQI KPLNHSDDLPHDRNGETIYCSCMIYLRPTNGAKLLMKWSEELQSQAWSESIRFKAN DQPAENLALNKTAHQVDLYLLSQVAFPTGGLYFKNEAWVQETKGKHIIVHNNYIIGYD RKMRFQDYGGLWLVDHHALESPLGKLE"

10826..12762

/gene="T7B11.4"

/note="encodes hypothetical protein; identical to T15B16.7, GenBank accession number AF104919; gene model last edited on 3 Mar 99"

/evidence=not_experimental

join(10826..10842,10929..11371,11559..11770,11843..12130,12260..12433,12526..12762)

/gene="T7B11.4"

/note="identical to T15B16.7, GenBank accession number AF104919; functional catalog ID=99"

/codon_start=1

/evidence=not_experimental

/product="hypothetical protein"

/protein_id="AAD22640.1"

/db_xref="GI:4558547"

/translation="MGTRHCYGLGYEMFNFWSSSVRFVWSSSETLFKIVITFRSLC ILGKVHTLLTFKVLTLFLKTELQCFVDCVYVAETGIVNIPPTKAQDGRPAGESGSK LRDEYILRGLNGLLFDKAYRVDGKGKDWKMDKPKLGLYGIAGADDVYNDGREGE ESAKLNAIQERTMRHIQKIVDDHERLTKLSESKKLEIKGNELAKPQVHNGTERMKL SEDLEQRQKEELHKKIIRLERIQDQVAIELEBQLKGQLNMKHMASDGDAAQVVKFV DIIFKDLVEKEAELADLNKFNQTLILRRRTNDELQEARKELVNMGELVYRKPFDVAM QOKYCOEDVEDRAVEVLQWEHYINDPDWHPYKRVKLENDREVEVIDDRDEKLRLK ADLGDGPNVATVKALLEINEYNPSGRYITTELWNFKEDKRALTEEGVTCLLDQWEKAK RKRGM"

```
gene
13384. .14244
/ gene="T7B11.5"
/ note="encodes RNasep-associated protein; identical to
T15B16.20, GenBank accession number AF104919; gene model
last edited on 3 Mar 99"
/ evidence=not_experimental
/ join(13384. .13454,13566. .13613,13697. .13757,13842. .13940,
14020. .14244)
/ gene="T7B11.5"
/ note="identical to T15B16.20, GenBank accession number
AF104919; functional catalog ID=98"
/ codon_start=1
/ evidence=not_experimental
/ product="putative RNasep-associated protein"
/ protein_id="AAD22641.1"
/ db_xref="GI:4558548"
/ translation="MRGVTPINPQKKIEADSNPVKESDYVEGERLTHLLDLIQRGIET
SKLSNVNSLPKIKWLKQIAIGINEVTVLRMNPNTSDQQNPKQLQVVLVADCK
PRMLTKHIPNLAASRVPLVYVRDNKRASRLGELVKLTALAIGIKARGNDLNLKK
QILPRDS"
complement(14442. .19647)
/ gene="T7B11.6"
/ note="encodes putative SecA-type chloroplast protein
transport factor; gene model last edited on 26 Mar 99"
/ evidence=not_experimental
complement(join(14442. .14579,14663. .14902,14991. .15108,
15196. .15263,15359. .15610,15686. .15787,15888. .15946,
16035. .16134,16215. .16361,16449. .16515,16710. .16885,
17107. .17274,17533. .17646,17718. .17831,17987. .18205,
18310. .18488,18577. .18684,18783. .18989,19082. .19154,
19231. .19647))
/ gene="T7B11.5"
/ note="functional catalog ID=06.04; functional catalog
ID=30.26"
/ codon_start=1
/ evidence=not_experimental
/ product="putative SecA-type chloroplast protein transport
factor"
/ protein_id="AAD22642.1"
/ db_xref="GI:4558549"
/ translation="MVSPLCDSQLLYHRPSISPTASQFVIADGIILRQNRLLSSSF
GTFKGNVTVKLGVSQSCSRKSTSVNASLGLLGIKFGSDNGESTRQQYASIVASV
NRLETEISALSDSELRERDALKQRAQKESMDSLPEAFVAVREASKRVGLRPPDV
QLIGGMVLHKGEIAEMRTGCKTLVAILPAYLNALSGRGVHVTVNDYLARRDCWVG
QVPRFLGLKVLGIQNMTPQQRKENYLCIDITYVTNSLGFDFLDNRNLATESVEELVLR
DFNYCVIDEVDLSILIDEARTPLIISGPAEKPSQYVKAAKIASAFERDIHTYDEKQK
TVLLTEQGYEDAEEHLDVLDYDPREQWASYVLNAIKAKELFLRDVNYIIRAKEVLIV
DEFTGRVMQRRWSDGLHQAWEAKGLPIQNESITLASISYQNFPLQFPKLCQMTGTA
STESAEFESIYKLVITVTPNPKMKDESDVVFKA VNGKRAVWEISRMHKTGRAV
LVGTTVSQSDLSQLLREAGITHEVLNAKPNVEREAEIVAQSGRLGAVTIATNMAG

Query Match      30.1%; Score 246; DB 8; Length 120185;
Best Local Similarity 100.0%; Pred. No. 3.1e-48;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 565 CAGGTGAGCGTGGGAGAAAGCAGAGAAGATGACTGGTGATATGCTTTCACCGTCTTAGG 624
|||||
Db 66752 CAGGTGAGCGTGGGAGAAAGCAGAGAAGATGACTGGTGATATGCTTTCACCGTCTTAGG 66811
|||||

QY 625 AACAAACAGAGCTCAAGAATGGTGTTCATGTTTCATTCGGTCTCTAGATTCTCGGAA 684
|||||
Db 66812 AACAAACAGAGCTCAAGAATGGTGTTCATGTTTCATTCGGTCTCTAGATTCTCGGAA 66871
|||||

QY 685 TAATAATGAATGGAGTCTGTGTTTGGTTCATGTTGAATCGATCAAGAAGTGTGTTTAAAC 744
|||||
Db 66872 TAATAATGAATGGAGTCTGTGTTTGGTTCATGTTGAATCGATCAAGAATGTGTTTAAAC 66931
|||||

QY 745 TGTACATGAATTATGCAGAAACATCTGTCCTGGTTCAGACATCGAAACTCTGTTCTTA 804
|||||
Db 66932 TGTACATGAATTATGCAGAAACATCTGTCCTGGTTCAGACATCGAAACTCTGTTCTTA 66991
|||||
QY 805 ATAAAA 810
|||||
```

```
Db 66992 ATAAAA 66997

RESULT 11
ATCHRIV5
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1 (bases 1 to 198220)
Spiegel, L.A., Huang, E.N., Nascimento, L.U., de la Bastide, M.,
Vil, D.M., Preston, R.R., Matero, A., Shah, R., O'Shaughnessy, A.,
Rodriguez, M., Shekher, M., Schutz, K., See, L.H., Swaby, I.,
Habermann, K., Dedhia, N.N., Mewes, H.W., Lemcke, K. and Mayer, K.F.X.
Unpublished
2 (bases 1 to 198220)
EU Arabidopsis sequencing, project.
Direct Submission
Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project
Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
E-mail: michael.bevan@bbsrc.ac.uk
Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/
this fragment has an overlap with ATCHRIV4 at the 5' end and an
overlap with ATCHRIV6 at the 3' end.
Location/Qualifiers
1. .198220
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
5603. .8671
/ gene="AT4g01810"
complement(join(5603. .5740,6062. .6178,6259. .8079,
8177. .8671))
/ gene="AT4g01810"
complement(join(5603. .5740,6062. .6178,6259. .8079,
8177. .8671))
/ gene="AT4g01810"
/ note="similarity to
contains EST gb:H36218, AA651596, T42768, AI997386.1,
AA395161"
/ codon_start=1
/ product="putative protein transport factor"
/ protein_id="CAB80674.1"
/ db_xref="GI:7268565"
/ db_xref="GOA:Q9SY11"
/ db_xref="SPTREMBL:Q9SY11"
/ translation="MANLPKSSVNPVGTLTPLENRPSPQDRTVPVHSPVVASPIP
PRFPQSFQPDQMSSPSMKSPSLSPANGIRTGSPIPRLSTPPGPPVPTPVKPAAVP
FRTSPATPQPMAYSSANSLPVSTPSFYNGSSVGSQSDLPDVVRMEEPIAADSPYVL
FSANKVLKOKKLANVASLGFCAIVSAGREISPGQIIQDRPHRCNLCGAYSNPYSSIL
IGSGQWCVICENMNGSKGYVASSKNELQNFPELSLPLVDYVQGNKRPFGVPASDS
RTSAPVVLVDECLDEPHLQSLHAFVDSLPGQTRGLIILYGRTVSIYDESEDSV
ASADVISGAKSPSABSMKALIYGTGVYLSPMHASKVAHEIFSSLRPYTLNVPESRD
RCLGTAVEAALAI IQPSAEMSRGVVRAGNSRIIVCAGGPITYGPGSVPHSMSPN
YPMKTAIKMENLGREAHRTVVDILCAGTCPLRVPILQPLAKASGGVVLVHDDF
GEAPGVDLQRAATRAAGSHGLLEVRCSDDILITQVIGPEEAHSETHETPKSDAALS
QMLSVETQSFSLSMENKRDIKSDHVFQFAFYSDVYQADSVRVTTFKLPVDSISA
YLQSVDEASAVLSIKRTLAKNQKDAVDMRATVDERIKDIALKFGSQVPSKLYSF
PKELSSLPPELLFHLRRGPLLGNIIGHEDERSVLNLFNASFDLSLRVAPRCLMHQE
GGTFEELPAYDLQSMQSDKAVILDHGTDFVFIWLGAELSADEVKSAAVLAACRTLAELT
```

EFRRPAPRILAFKEARFPQIRTLTTEQRMKLKSSFLFDEASFCWMRSLKVVPEPEPR
" complement (5603. .5740)
/gene="AT4g01810"
/number=1
intron complement (5741. .6061)
/gene="AT4g01810"
/number=1
exon complement (6062. .6178)
/gene="AT4g01810"
/number=2
intron complement (6179. .6258)
/gene="AT4g01810"
/number=2
exon complement (6259. .8079)
/gene="AT4g01810"
/number=3
intron complement (8080. .8176)
/gene="AT4g01810"
/number=3
exon complement (8177. .8671)
/gene="AT4g01810"
/number=4
gene 9603. .14198
/gene="AT4g01820"
complement (join(9603. .9956,10053. .10349,10502. .10765,
10835. .11101,11218. .12029,12106. .12905,12987. .13225,
13293. .13514,13592. .13767,13850. .13904,13995. .14198))
/gene="AT4g01820"
complement (join(9603. .9956,10053. .10349,10502. .10765,
10835. .11101,11218. .12029,12106. .12905,12987. .13225,
13293. .13514,13592. .13767,13850. .13904,13995. .14198))
/gene="AT4g01820"
/note="evidence: experimental, likely to be involved in
lipophilic transport
contains EST gb:AI996312.1, AI995010.1, AI994333.1,
R90081, AA605450"
/codon_start=1
/product="P-glycoprotein-like protein pgp3"
/protein_id="CAB80675.1"
/db_xref="GI:7268566"
/db_xref="GOA:Q9SYI2"
/db_xref="SPTREMBL:Q9SYI2"
/translation="MEETKTVPFVKLFSPSDSTVLLMIVGSGIAGNGVGFPLMTL
LFGLIDISIGQNSKDIIVEISKVCLKFLVYGLGLTGAAPFLQVACMMITGERQAARI
RSLYLTLRLQDIGFFDVETSTGEVVGSRMGSDTVLILEAMGEKVKFQIQLIATFVGGE
VLAFFVKGLLTLVMLVSIPLLAIAAGAAMP IIVTRASSREQAAYAKASTVVEQTLGSIR
TVASFTGEKQAMKSYREFINLAYRSVKQFSGMLGLGVVFFVFFCSYALAIWFGGEM
ILKKGVTGGEVNVMTVVASMSLGQTPCLTAFAGKAAAYKMPETIERKPSIDAF
DLNGKVLIEDIRGEIELRDVCFSPARPMEVEFGFSLIPSGATAALVGESGSGKSSV
LSLIERFYPDPSSGVLIDGNLKEFLKWIRGKIGLVSPVLPFSSSIMENIGYKEN
ATVEEIQAAKLANAANFIDKLPRGLETLVGEHGTQSGGQKQRIATARAILKDPRIIL
LLDEATSALDAESERVVQEQALDRVMSRTTVIVAHRLSTVRNADMFAVHRGKIVEEG
SHSELKDHGAYAQILRLQIKKEPKRLESSNELRDRSINRGSRNIRTVHDDDSV
SVLGLLGRQENTEISREQSRNVSIIRIAALNKPETTILILGTLGAVNGTIFPIPGIL
FAKVIEAFKPPHDMKDRSFRFSMIFVLIGVASLIVYPMHTYLFVAVAGRLLIQIRVM
CFEKVVMVEVGFDDPENSSGTIGSRLSADAALIKTLVGDSLVSLSVKNAAAASGLII
AFTASWKLAVIILVMIPLIGINGYLQIKFIKGTADAKAKYEEASQVANDAVGSIRTV
ASFCAEKVMEMYKKRCEDTIKSGIKQGLISGVGFGISFFVLVSIVASCFFVVGARLVK
AGRTNENDVFQVFLATMTAIGISQASSFAPDSSKAKGAAASIFGIIDGKSMIDSRDE
SGLVLENKGDIELCHISFTYQTRPDVQIFRDLCAIRAGQTVLAVGESGSGKSTVIS
LLQRFYDPSCHITLDRVELKKLQKLVQRMGLVQEPVLENDTIRSNIAVYGGKDE
ASEAEIIAARELANAHGFISIIQQGYDTVVGERGIQLSGGQKQRAIARAIKPEPKIL
LLDEATSALDAESERVVQDALDRVMNRITTVVAHRLSTIKNADVIAVKNGVIVEKG
THETLINIEGGVYASLVQLHISASS"
complement (9603. .9956)
/gene="AT4g01820"
/number=1
intron complement (9957. .10052)
/gene="AT4g01820"
/number=1
exon complement (10053. .10349)
/gene="AT4g01820"

intron /number=2
complement (10350. .10501)
/gene="AT4g01820"
exon /number=2
complement (10502. .10765)
/gene="AT4g01820"
/number=3
complement (10766. .10834)
/gene="AT4g01820"
/number=3
complement (10835. .11101)
/gene="AT4g01820"
/number=4
complement (11102. .11217)
/gene="AT4g01820"
/number=4
complement (11218. .12029)
/gene="AT4g01820"
/number=5
complement (12030. .12105)
/gene="AT4g01820"
/number=5
complement (12106. .12905)
/gene="AT4g01820"
/number=6
complement (12906. .12986)
/gene="AT4g01820"
/number=6
complement (12987. .13225)
/gene="AT4g01820"
/number=7
complement (13226. .13292)
/gene="AT4g01820"
/number=7
complement (13293. .13514)
/gene="AT4g01820"
/number=8
complement (13515. .13591)
/gene="AT4g01820"
/number=8
complement (13592. .13767)
/gene="AT4g01820"
/number=9
complement (13768. .13849)
/gene="AT4g01820"
/number=9
complement (13850. .13904)
/gene="AT4g01820"
/number=10
complement (13905. .13994)
/gene="AT4g01820"
/number=10
complement (13995. .14198)
/gene="AT4g01820"
/number=11
14552. .19316
/gene="AT4g01830"
complement (join(14552. .14905,14994. .15290,15475. .15738,
15809. .16075,16311. .17110,17180. .17979,18068. .18306,
Query Match 30.1%; Score 246; DB 8; Length 198220;
Best Local Similarity 100.0%; Pred. No. 3e-48;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 565 CAGGTGACCGTGGGAGAAAGCAGAGAAGATGACTGGTGATGATGCTTTCACCGTCTTAGG 624
Db 52105 CAGGTGACCGTGGGAGAAAGCAGAGAAGATGACTGGTGATGATGCTTTCACCGTCTTAGG 52164
Qy 625 AACAAACAGAGCTCAAGATGCTTTTTTTTTTTCATTCGGTCTCTAGATTCGCGAA 684
Db 52165 AACAAACAGAGCTCAAGATGCTTTTTTTTTTTCATTCGGTCTCTAGATTCGCGAA 52224
Qy 685 TAATAATGAATGGAGTCTCTGTGTTTGGTTTTCATGTTGAATCGATCAAGATGTGTTTTTAAC 744


```
QY 332 TGTTACTGTTTCTGATGTCAGAGGGTTTGGTGCACRAGGAGGTTCTACCGAGAGACACGG 391
Db 431 CGTAACTGTTCTGATGTTAGAGGTTTGGAGTTTCAGGCTGGATCTGCAGAACGGCAAGC 490
QY 392 TGGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAA 451
Db 491 AGGCTCCGAGTTTCTTAAAGACAAATTTGTGTCAAAAGTAAAGATGGAGATTTGTGTATC 550
QY 452 GAAACACCAAGTGAATCTGTAATCAACACAATAATTGAAGGAGCAAGGACAGGAGAT 511
Db 551 TAAAGATCAGGTAGAACGAGTAATTGATGCAATCATTTGATGAGGCAAGAACTGGAGAAAT 610
QY 512 TGGTGATGCGCAAGATTTTGTGTTTGGCTGTGTCTAGATGTCATAGAGTTAGGACAGGTGA 571
Db 611 TGGAGATGGCAAAATATTTGTGTTCCAGTTGCAGATGTCATTCGTGTGAGAACAGGTGA 670
QY 572 GCGTGGGGAAGACGAGAGAAGATGACTGGTG 603
Db 671 GCGGGACTTGAAGCCAGAGAGAAATGGCTGGTG 702

RESULT 14
AK068407
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) cDNA clone:J013149B08, full
insert sequence.
AK068407
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
1
The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-Length cDNA Project Team:
Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group: Otomo,Y., Murakami,K.,
Iida,Y., Sugano,S., Fujimura,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Yoshino,M. and Hayashizaki,Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 {5631}, 376-379 (2003)
22752273
12869764
2 (bases 1 to 902)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K.,
Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,
Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K.,
```

```
SAKAZUME,N., SANO,H., SASAKI,D., SATO,K., SATOH,K., SHIBATA,K.,
SHINAGAWA,A., SHIRAKI,T., SHISHIKI,T., SOGABE,Y., SUGANO,S.,
SUGIYAMA,A., SUZUKI,K., TAGAMI,M., TAGAMI,M., TAGAMI-TAKEDA,Y.,
TAGAWA,A., TAKAHASHI,F., TAKAKU-AKAHIRA,S., TANAKA,T., TOMARU,A.,
TOYA,T., TSUNODA,Y., UEDA,M., WAKI,K., XIE,Q., YAHAGI,W.,
YAMADA,H., YAMAMOTO,M., YASUNISHI,A., YAZAKI,J., YOKOMIZO,S. and
YOSHIMURA,A.
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica
rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K.,
Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y.,
Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N.,
Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,
Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,
Yasunishi,A. and Hayashizaki,Y.
Location/Qualifiers
1..902
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J013149B08"

FEATURES
Source
Query Match 24.4%; Score 199.4; DB 8; Length 902;
Best Local Similarity 71.3%; Pred. NC. 6.4e-37;
Matches 263; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

ORIGIN
QY 240 CCAGACTCGAAATTTTACRAGGTGGAAGCAATTTGTCAGACCATCGAGAACATCCAGCAAGTT 299
Db 283 CCGGAGTCGGAGTTCTACRAGGTGGAGGCAATCTTGAGGCCATGAGGGTGCCTTATGTG 342
QY 300 TCATCGGCTTTACTGAAATTCGGGATTCGAGGTGTTACTGTTTCTGATGTGAGAGGGTTT 359
Db 343 TCATCGGCTTTGCTGCAATGGGATCAGAGCGGTGACGGTGTCCGACGTGCCGGGTTTC 402
QY 360 GGTGCACAAGGAGGTTCTACCGAGAGACACGGTGGCTCTGAGTTCTCGGAAGACAAATTT 419
Db 403 GCGGCACAGGCGGGTCAACTGAGAGGCATGAGGGTTCAGAAATTTGCAGAAGATACATTT 462
QY 420 GTTGCTAAAGTTAAGATGGAATCGTTGTTAAGAAAGACCAAGTGAATCTGTAATCAAC 479
Db 463 ATTGATAAGTTAAGATGGAATAGTGTGTCCAGGATCAGGTTGAAGCTGTTGTTGAC 522
QY 480 ACAATAATTGAAGGAGCAAGGACAGGAGCATTTGGTGATGGCAAGATTTTGTGTTTCCT 539
Db 523 AAGATAATTGAAAAGGCAAGAACAGGAGAAATTTGGTGATGGAAAATATTTTGTATACCC 582
```

Qy 540 GTGTGACATGTCATAGAGTTAGGACAGCGTGAGCGTGCGGAGAAAAGCAGAGAGATGACT 599
|||||
Db 583 GTGTGGACGTGATCAGATACGACCGCGGAACGAGCGGAGCGGAGCGGATGGCC 642
|||||
Qy 600 GGTGATATG 608
|||
Db 643 GGAGGGCTG 651
|||

RESULT 15
AK099152
LOCUS
DEFINITION
AK099152 902 bp mRNA linear PLN 24-JUL-2003
Oryza sativa (japonica cultivar-group) cDNA clone:J023070H02, full
insert sequence.
AK099152
AK099152.1 GI:32984361
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE
AUTHORS
1 The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-Length cDNA Project Team:
Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
Kishimoto,M., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group: Ootomo,Y., Murakami,K.,
Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Yoshino,M. and Hayashizaki,Y.

TITLE
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)
22752273

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
2 (bases 1 to 902)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashi,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
Kodama,T., Kojima,K., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
Koya,S., Kurihara,C., Kurosaki,T., Matsuyama,T., Miura,J., Miyazaki,A.,
Masuda,H., Matsubara,K., Murata,M., Nagata,T., Nakamura,M.,
Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K.,
Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,
Osato,N., Ota,Y., Ootomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K.,
Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
Yoshimura,A.

TITLE
JOURNAL
Direct Submission
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp).

COMMENT
Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica
rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K.,
Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
Yamamoto,M.

PAIS Genome Sequencing & Analysis Group: Ootomo,Y., Iida,Y.,
Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N.,
Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,
Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,
Yasunishi,A. and Hayashizaki,Y.

FEATURES
source

1..902
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023070H02"

ORIGIN

Query Match 24.4%; Score 199.4; DB 8; Length 902;
Best Local Similarity 71.3%; Pred. No. 6.4e-37;
Matches 263; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

Qy 240 CCAGACTCGAAATTTTACAAGGTGGAGCAATGTGCAGACCATGGAGATCCAGCAAGTT 299
|||
Db 283 CCGGAGTCGAGTTCTACAAGGTGGAGCAATCTTGAGCCATGGAGGTGCTTATGTG 342
|||
Qy 300 TCATCGGCTTTACTGAAATCGGATTCAGGTGTACTGTCTCTGATGTGAGGGTTT 359
|||
Db 343 TCATCGGCTTTCTGCAATGGGATCAGAGCGGTGACGGTGTCCGACGTCGCGGGTTTC 402
|||
Qy 360 GGTGCACAAGGAGGTTCTACCGAGAGACACGGTGGCTCTGAGTTCTCGGAAGACAAATTT 419
|||
Db 403 GCGGCACAGCGCGGTCAACTGAGAGGCGATGAAGGTTCAGAAATTTGCAGAAGATACATTT 462
|||
Qy 420 GTTGCTAAAGTTAAGATGGAATCGTTGTTAAGAAAGACCAAGTGGAAATCTCTAATCAAC 479
|||
Db 463 ATTGATAAAGTTAAGATGGAATAGTGGTGTCCAGGATCAGGTTGAAGCTGTGTGAC 522
|||
Qy 480 ACAATAATTGAAGGACCAAGGACAGGAGAGATTGGTGTATGGCAAGATTTTTCCTTGCCT 539
|||
Db 523 AAGATAATTGAAAAGCAAGAACAGGAGAAATTTGGTGTGGAATAATTTTGTATACCC 582
|||
Qy 540 GTGTGAGATGTCATAAGAGTTAGGACAGGTGCGGTGGGAGAAAGCAGAGAGATGACT 599
|||
Db 583 GTGTGCGACGTGATCAGATACGACCGCGGAACAGGGGAGCGGAGGAGGATGGCC 642
|||
Qy 600 GGTGATATG 608
|||
Db 643 GGAGGGCTG 651
|||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 13:28:59 ; Search time 284.934 Seconds
(without alignments)
12180.972 Million cell updates/sec

Title: US-09-756-541-13
Perfect score: 817
Sequence: 1 CTGAAAGTTGTGTTAAAAA.....GTTCTTAATAAAAAAAA 817

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	817	100.0	817	AAF58581	Aaf58581 Arabidops
2	591	72.3	591	ABZ14328	Abz14328 Arabidops
3	588	72.0	588	AAF58583	Aaf58583 Arabidops
4	286.4	35.1	897	AAF58582	Aaf58582 Ricinus c
5	285	34.9	594	AAF58584	Aaf58584 Ricinus c
6	92.8	11.4	96109	AAF28548	Aaf28548 Genomic f
7	85.8	10.5	375	ADA30024	Ada30024 DNA encod
8	84.6	10.4	110000	AAV21209_01	Continuation (2 of
9	84.4	10.3	110000	AAV21209_14	Continuation (15 o
10	81.6	10.0	110000	ACF67367_52	Continuation (53 o
11	81.6	10.0	110000	ACF65387_3	Continuation (4 of
12	81.4	10.0	339	ACF72027	Acf72027 Photorhab
13	79.8	9.8	110000	AAT42063_03	Continuation (4 of
14	77	9.4	363	ABZ40022	Abz40022 N. gonorr
15	73	8.9	16526	AAA81472	Aaa81472 N. mening
16	73	8.9	110000	AAA81489_6	Continuation (7 of
17	73	8.9	172325	AAF21613	Aaf21613 Neisseria
18	73	8.9	349980	AAF21612	Aaf21612 Neisseria
19	68.8	8.4	336	AAH68445	Aah68445 C glutami
20	68.8	8.4	417	AAH67222	Aah67222 C glutami
21	68.8	8.4	349980	AAH68531	Aah68531 C glutami
22	67.2	8.2	1164	AAS88441	Aas88441 DNA encod
23	67.2	8.2	11095	AAS46256	Aas46256 DNA encod

24	64.6	7.9	663	7	ACF65588	Acf65588 Photorhab
25	64.6	7.9	110000	7	ACF67367_46	Continuation (47 o
26	64.6	7.9	110000	7	ACF65388_01	Continuation (2 of
27	61	7.5	339	7	ACF71482	Acf71482 Photorhab
28	60.4	7.4	110000	6	ABQ69245_15	Continuation (16 o
29	60.4	7.4	110000	6	ABQ67195_1	Continuation (2 of
30	60.4	7.4	110000	6	ABQ67195_2	Continuation (3 of
31	59.6	7.3	290	4	AAF71296	Aaf71296 Corynebact
32	57.6	7.1	110000	6	ABA90521_16	Continuation (17 o
33	52.4	6.4	2751	6	ABQ70750	Abq70750 Listeria
34	50.8	6.2	110000	6	ABA03041_15	Continuation (16 o
35	49.6	6.1	110000	4	AAI99682_32	Continuation (33 o
36	49.6	6.1	110000	4	AAI99683_32	Continuation (33 o
37	47	5.8	342	6	ABN18985	Abn18985 Human ORF
38	45.8	5.6	9733	6	ABL32682	Ab132682 Human imm
39	44.4	5.4	174	7	ABZ39932	Abz39932 N. gonorr
40	43.6	5.3	2000	7	ADA71938	Ada71938 Rice gene
41	43.4	5.3	3858	5	AAS89065	Aas89065 DNA encod
42	43.4	5.3	3858	5	AAS90406	Aas90406 DNA encod
43	43.4	5.3	3858	5	AAS85755	Aas85755 DNA encod
44	43.4	5.3	3858	5	AAS92604	Aas92604 DNA encod
45	42.8	5.2	1970	3	AAZ99179	Aaz99179 T. hydrot

ALIGNMENTS

RESULT 1
AAF58581
ID AAF58581 standard; cDNA; 817 BP.
XX
AC AAF58581;
XX
DT 23-APR-2001 (first entry)
XX
DE Arabidopsis thaliana P-PII cDNA.
XX
KW Arabidopsis thaliana; PII; plant nitrogen regulatory gene; P-PII;
nitrogen assimilation; transgenic plant; herbicide screening; ss.
XX
OS Arabidopsis thaliana.
XX
PN US6177275-B1.
XX
PD 23-JAN-2001.
XX
PF 23-JUL-1997; 97US-00899330.
XX
PR 24-JUL-1996; 96US-0022328P.
XX
PA (UUNY) UNIV NEW YORK STATE.
XX
PI Coruzzi GM, Lam H, Hsieh M;
XX
DR WPI; 2001-158572/16.
DR P-PSDB; AAB69495.
XX
PT Novel P-PII genes capable of regulating plant nitrogen assimilation,
useful for transgenic plant production, and as probes for isolating
additional genomic clones having P-PII gene promoters.
XX
PS Claim 2; Fig 12; 35pp; English.
XX
CC The present sequence encodes a nitrogen regulatory PII protein. Novel
plant PII (also called P-PII) nucleotide sequences have been isolated.
CC They are useful for regulating nitrogen assimilation in plants, and in
transgenic plant production. They are also used to engineer organisms
that overexpress wild-type or mutant P-PII regulatory proteins. P-PII
proteins are useful for in vitro screening of herbicides. P-PII
nucleotides may be used as probes for isolating additional genomic clones
with the promoters of P-PII genes. P-PII promoters are light- and/or
sucrose-inducible, and are suitable for genetic engineering of plants

SQ	Sequence 817 BP; 242 A; 137 C; 188 G; 250 T; 0 U; 0 Other;	
	Query Match 100.0%; Score 817; DB 4; Length 817;	
	Best Local Similarity 100.0%; Pred. No. 3.8e-220;	
	Matches 817; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 CTGAAGTGTGTTAAATAAACTAGAAATCATGGCGGCTCAATGACGAAACCCATCT 60	
Db		
QY	1 CTGAAGTGTGTTAAATAAACTAGAAATCATGGCGGCTCAATGACGAAACCCATCT 60	
Db		
QY	61 CAATAACTTCTCTCGGTTTCTATTCTGATCGAAGAACATTTGCTTCTCTGATTGCAATT 120	
Db		
QY	61 CAATAACTTCTCTCGGTTTCTATTCTGATCGAAGAACATTTGCTTCTCTGATTGCAATT 120	
Db		
QY	121 CGATTGTTCTGGATTCTGATTCAGACATTTCCGACCAATCTTGCTCGAATGGTCACAAAGTCAC 180	
Db		
QY	181 CGAGTAATAACAGTCGTGTTTACCTGTCGTTAGTGCCCAATACTCTGATTATATTC 240	
Db		
QY	241 CAGACTCGAAATTTTACAAGGTGGAAGCAATTTGTGAGACCATGAGAAATCCAGCAAGTTT 300	
Db		
QY	301 CATCGGCTTTACTGAAATCGGATTCGAGTTCTGAGTTTCTGATGAGAGGTTT 360	
Db		
QY	361 GTGCACAAGGAGTTTCTACCGAGAGACACGCTGCTCTGAGTTCTCGAAGACAAATTTG 420	
Db		
QY	361 GTGCACAAGGAGTTTCTACCGAGAGACACGCTGCTCTGAGTTCTCGAAGACAAATTTG 420	
Db		
QY	421 TTGCTAAAGTTAAGATGGAATCGTTTGAAGAACCAAGTGGATCTGTAATCAACA 480	
Db		
QY	481 CAATAATTGAAGAGCAAGGACAGGAGATTTGTTGATGGCAAGATTTTGTGCTG 540	
Db		
QY	541 TGTCAGATGTCATAAGAGTTAGGACAGGTCAGGCTGGGAGAAAGCAGAGAGATGACTG 600	
Db		
QY	501 GTGATATGCTTTACCGTCTTAGGAACAACACAGAGCTCAAGAAATGTTTTTTTTC 660	
Db		
QY	501 GTGATATGCTTTACCGTCTTAGGAACAACACAGAGCTCAAGAAATGTTTTTTTTC 660	
Db		
QY	661 ATTTGGTCTCTAGATTCTGCGAATAATAATGAATGAGTCTGTGTTTGGTTTCATGTTG 720	
Db		
QY	661 ATTTGGTCTCTAGATTCTGCGAATAATAATGAATGAGTCTGTGTTTGGTTTCATGTTG 720	
Db		
QY	721 AATCGATCAAGATGTGTTTTTAACTGTACATGAATTATGCAGAAACATCTGCTGGTTC 780	
Db		
QY	781 TCAGACATCGAAACTCTGTTCTTAATAAAAAA 817	
Db		
QY	781 TCAGACATCGAAACTCTGTTCTTAATAAAAAA 817	
Db		
RESULT 2		
ABZ14328		
ID	ABZ14328 standard; DNA; 591 BP.	
XX		
AC	ABZ14328;	
XX		
DT	21-JAN-2003 (first entry)	
XX		
DE	Arabidopsis thaliana stress regulated gene SEQ ID NO 2133.	

QY	453	AAAGACCAAGTGGAAATCTGTAATCAACACAAATAATTGAAGGAGCAAGACAGGAGAGATT	512
Db	421	AAAGACCAAGTGGAAATCTGTAATCAACACAAATAATTGAAGGAGCAAGACAGGAGAGATT	480
QY	513	GGTGATGGCAAGATTTTGTGTTTGCCTGTGTGATGTCATAAGAGTTAGGACAGGTTGAG	572
Db	481	GGTGATGGCAAGATTTTGTGTTTGCCTGTGTGATGTCATAAGAGTTAGGACAGGTTGAG	540
QY	573	CGTGGGGAGAAAGCAGAGAAGATGACTGGTGATATGCTTTTCACCGTCCTTAG	623
Db	541	CGTGGGGAGAAAGCAGAGAAGATGACTGGTGATATGCTTTTCACCGTCCTTAG	591

RESULT 3

AAF58583
ID AAF58583 standard: cDNA: 588 bp.

AC AAF58583;

DT 23-APR-2001 (first entry)

DE *Arabidopsis thaliana* P-PII cDNA fragment.

Arabidopsis thaliana; PII; plant nitrogen regulatory gene; P-II; KW
nitrogen assimilation; transgenic plant; herbicide screening; ss. KW

OS Arabidopsis thaliana.

PN US6177275-B1.

23-JAN-2001.

23-JUL-1997: 97US-00899330.

PR 24-JUL-1996: 96US-0022328P.

PA (UYN) UNIV NEW YORK STATE.

PI Coruzzi GM, Lam H, Hsieh M;

WPI: 2001-158572/26.

DR P-PSDB; AAB69495.

Novel P-II genes capable of regulating plant nitrogen assimilation, useful for transgenic plant production, and as probes for isolating additional genomic clones having P-II gene promoters.

PS Claim 2: Col 37-38; 35pp: English.

The present sequence encodes a nitrogen regulatory PII protein. Novel plant PII (also called P-PII) nucleotide sequences have been isolated. They are useful for regulating nitrogen assimilation in plants, and in transgenic plant production. They are also used to engineer organisms that overexpress wild-type or mutant P-PII regulatory proteins. P-PII proteins are useful for in vitro screening of herbicides. P-PII nucleotides may be used as probes for isolating additional genomic clones with the promoters of P-PII genes. P-PII promoters are light- and/or sucrose-inducible, and are suitable for genetic engineering of plants

Sequence 588 BP; 164 A; 105 C; 147 G; 172 T; 0 U; 0 Other;

```

Query Match          72.0%; Score 588; DB 4; Length 588;
Best Local Similarity 100.0%; Pred. No. 1.4e-155;
Matches 588: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 33 ATGGCGGGCTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGA 92

Dh 1 ATGCGCGGGCTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGA 60

93 AAGACATTGCTTTCTCTGATTGCATTTTCGATTTGTTCTGGAATTCAGACATTCGGACCA 152

61 AAGACATTGCTTTCTCTGATTGCATTTTCGATTTGTTCTGGAATTCAGACATTCGGACCA 120

QY	153	TCTTGCCTCGATTTGGTCACAAAGTCACCGAGTAATAACAGTCGTGTTTTACCTGTCGTT	215
Db	121	TCTTGCCTCGATTTGGTCACAAAGTCACCGAGTAATAACAGTCGTGTTTTACCTGTCGTT	180
QY	213	AGTGCCCAATATCTTCTGATTTATATCCAGACTCGAAATTTTACAAGTGGAAGCAATT	272
Db	181	AGTGCCCAATATCTTCTGATTTATATCCAGACTCGAAATTTTACAAGTGGAAGCAATT	240
QY	273	GTGAGACCATGGAGATCCAGCAAGTTTTCATCGGCTTTTACTGAAATTCGGGATTCGAGGT	332
Db	241	GTGAGACCATGGAGATCCAGCAAGTTTTCATCGGCTTTTACTGAAATTCGGGATTCGAGGT	300
QY	333	GTACTGTGTTCTCGAAGACAAAATTTGTTGTCACAAAGGAGTTCTTACCGAGAGACACGGT	392
Db	301	GTACTGTGTTCTCGAAGACAAAATTTGTTGTCACAAAGGAGTTCTTACCGAGAGACACGGT	360
QY	393	GGCTCTGAGTTCTCGGAAGACAAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAG	452
Db	361	GGCTCTGAGTTCTCGGAAGACAAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAG	420
QY	453	AAAGACCAAGTGGAAATCTGTAATCAACACAATTAATTGAAGGACCAAGGACGGAGAGATT	512
Db	421	AAAGACCAAGTGGAAATCTGTAATCAACACAATTAATTGAAGGACCAAGGACGGAGAGATT	480
QY	513	GGTGATGGCAAGATTTTGTGTTTGCCCTGTGTGATGTCATTAAGATTAGGACAGGTGAG	572
Db	481	GGTGATGGCAAGATTTTGTGTTTGCCCTGTGTGATGTCATTAAGATTAGGACAGGTGAG	540
QY	573	CGTGGGAGAAAGCAGAGAAGATGACTGGTGATATGCTTTTCAACCGTCT	620
Db	541	CGTGGGAGAAAGCAGAGAAGATGACTGGTGATATGCTTTTCAACCGTCT	588

RESULT 4

AAF58582

ID: AAF58582 standard; cDNA; 897 BP.

AC AAF58582

23-APR-2001	(first entry)
-------------	---------------

DE Rictinus communis P-PII cdNA.

KW Castor bean; PII; plant nitrogen regulatory gene; P-II;
KW nitrogen assimilation; transgenic plant; herbicide screening; ss.

Ricinus communis.

PN US6177275-B1.

23--JAN-2001.

23--JUL-1997: 97US-00899330.

24-JUL-1996: 96US-0022328P.

XX PA (JFNY) UNIV NEW YORK STATE.

Corruzzi GM- Lam H. Hsieh M-

DR WPT: 2001-158572/16.

DR P-PSDB: AAB69496.

Novel P-PII genes capable of regulating plant nitrogen assimilation, useful for transgenic plant production, and as probes for isolating additional genomic clones having P-PII gene promoters.

PS Claim 2: Fig 13: 35pp: English.

The present sequence encodes a nitrogen regulatory PII protein. Novel plant PII (also called P-PII) nucleotide sequences have been isolated. They are useful for regulating nitrogen assimilation in plants, and in transgenic plant production. They are also used to engineer organisms

RESULT 6
AAF28548
ID AAF28548 standard; DNA; 96109 BP.

XX AC AAF28548;
XX DT 04-APR-2001 (first entry)
XX DE Genomic fragment #35.

XX KW Genomic library; bacteria; human upper airway; otitis media; sinusitis;
XX KW bronchopulmonary; endocarditis; meningitis; ss.

XX OS Moraxella catarrhalis.
XX PN WO200078968-A2.

XX PD 28-DEC-2000.

XX PF 16-JUN-2000; 2000WO-US016649.

XX PR 18-JUN-1999; 99US-0140121P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Lagace RE, Patterson C, Berg KL;

XX DR WPI; 2001-041427/05.

XX PT Genomic library for identifying diagnostic and therapeutic compositions,
XX PT and for identifying virulence factors, regulatory elements and drug
XX PT targets, comprises Moraxella catarrhalis nucleic acids.

XX PS Claim 1; Page 345-368; 545pp; English.

XX CC The present invention relates to a Moraxella catarrhalis genomic library
XX CC comprising a combination of 41 nucleic acid molecules (see AAF28514-
XX CC AAF28554). The library has a number of uses described in the
XX CC specification e.g. is useful for identifying diagnostic and therapeutic
XX CC compositions. M. catarrhalis (Branhamella catarrhalis) is a large
XX CC aerobic, gram-negative diplococcus, normally found among the bacterial
XX CC flora of human upper airways. M. catarrhalis is known to cause acute,
XX CC localised infections such as otitis media, sinusitis and bronchopulmonary
XX CC infection and life-threatening, systemic diseases including endocarditis
XX CC and meningitis

XX SQ Sequence 96109 BP; 28783 A; 18910 C; 20341 G; 28075 T; 0 U; 0 Other;

Query Match 11.4%; Score 92.8; DB 4; Length 96109;
Best Local Similarity 57.8%; Pred. No. 7.1e-15;
Matches 185; Conservative 0; Mismatches 132; Indels 3; Gaps 1;

QY 267 GCAATTTCAGACCATGGAGAATCCAGCAAGTTTTCATCGGCTTTACTGAAATCGSGATT 326
DB 12953 GCAATTATCAAGCCGTTTAAACTCGATGATGCGGAGAAGCACTCTCAGAAATGGCGTC 13012

QY 327 CGAGGTGTTACTGTTCTGATGTGAGAGGGTTTGGTGCAAGGAGGTTTCTACCGAGAGA 386
DB 13013 AATGGTATCACCGTCACTGAAGTCAAGGCTTGGTCGCCAAAAGGTCATACCGAGATG 13072

QY 387 CACGGTGGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGTGAATCGTT 446
DB 13073 TATCGTGGGCGGAATATGTGGTTGAT---TTTTCACAAAATTAATTAATGAGATAGCA 13129

QY 447 GTTAAGAAAGACCAAGTGGAAATCTGTAATCAACAATAATTGAAGGAGGACAGGA 506
DB 13130 TGTCGTGATGAGATGGTTGATTCAAATTAATTGAGTCAATCATTAAAGTTGCAATACAGGT 13189

QY 507 GAGATTGGTGATGGCAAGATTTTGTGTTTGGCTGCTGTCAGATGTCATAAGAGTTAGGACA 566
DB 13190 AAAATTGGTGATGGTAAGATTTTGTGTTAGTCCGCTTGAGCGGTGTCATTCGCATTGCAACT 13249

QY 567 GGTGACCGTGGGGAGAAAGC 586

DB 13250 GCGGAATTGATGAAAGTGC 13269

RESULT 7
ADA30024

ID ADA30024 standard; DNA; 375 BP.

XX AC ADA30024;

XX DT 20-NOV-2003 (first entry)

XX DE DNA encoding Acinetobacter baumannii protein #1311.

XX KW ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;
XX KW vaccine; plant biocontrol agent.

XX OS Acinetobacter baumannii.

XX PN US6562958-B1.

XX PD 13-MAY-2003.

XX PF 04-JUN-1999; 99US-00328352.

XX PR 09-JUN-1998; 98US-0088701P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX PI Breton G, Bush D;

XX DR WPI; 2003-576092/54.

XX DR P-PSDB; ADA34150.

XX PT New Acinetobacter baumanii proteins and nucleic acids, useful as reagents
XX PT for diagnosing a bacterial disease, as components of antibacterial
XX PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
XX PT plants.

XX PS Example; SEQ ID NO 1311; 328pp; English.

XX CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
XX CC The A. baumannii nucleic acids and polypeptides are useful as reagents
XX CC for diagnosing a bacterial disease, as components of antibacterial
XX CC vaccines, as targets for antibacterial drugs, to detect the presence of
XX CC A. baumannii and other Acinetobacter species in a sample, in screening
XX CC compounds for the ability to interfere with the A. baumannii life cycle
XX CC or to inhibit A. baumannii infection, and as biocontrol agents for
XX CC plants. The present sequence represents DNA encoding an A. baumannii
XX CC protein.

XX SQ Sequence 375 BP; 122 A; 64 C; 93 G; 96 T; 0 U; 0 Other;

Query Match 10.5%; Score 85.8; DB 8; Length 375;
Best Local Similarity 57.4%; Pred. No. 6.5e-14;
Matches 175; Conservative 0; Mismatches 127; Indels 3; Gaps 1;

QY 267 GCAATTGTCAGACCATGGAGAATCCAGCAAGTTTTCATCGGCTTTACTGAAATCGSGATT 326
DB 52 GCAATTGTAAACCGTTTAAATTGGATGATGCGTGAAGCACTCTCTGACATTGGTGTGTA 111

QY 327 CGAGGTGTTACTGTTTCTGATGTGAGAGGGTTTGGTGCAAGGAGGTTCTACCGAGAGA 386
DB 112 CAAGGGATTACCGTAACTGAAGTTAAAGGTTTGGTCGTCAAAAGGACATACAGAACTT 171

QY 387 CACGGTGGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTT 446
DB 172 TACCGCGGCGCTGAGTATGTGGTTGAT---TTCTTACCTAAAGTAAATCGAAATGCG 228

QY 447 GTTAAGAAAGACCAAGTGGAAATCTGTAATCAACACAATAATTGAAGGAGCAAGCAGGA 506
DB 229 ATTAGTGATGAATGTCGACCGGTAATTGAGTCAATTACACGTGTGGCAAGCACTGGA 288

WP	ACF67367_28	2800001	2910000	
WP	ACF67367_29	2900001	3010000	
WP	ACF67367_30	3000001	3110000	
WP	ACF67367_31	3100001	3210000	
WP	ACF67367_32	3200001	3310000	
WP	ACF67367_33	3300001	3410000	
WP	ACF67367_34	3400001	3510000	
WP	ACF67367_35	3500001	3610000	
WP	ACF67367_36	3600001	3710000	
WP	ACF67367_37	3700001	3810000	
WP	ACF67367_38	3800001	3910000	
WP	ACF67367_39	3900001	4010000	
WP	ACF67367_40	4000001	4110000	
WP	ACF67367_41	4100001	4210000	
WP	ACF67367_42	4200001	4310000	
WP	ACF67367_43	4300001	4410000	
WP	ACF67367_44	4400001	4510000	
WP	ACF67367_45	4500001	4610000	
WP	ACF67367_46	4600001	4710000	
WP	ACF67367_47	4700001	4810000	
WP	ACF67367_48	4800001	4910000	
WP	ACF67367_49	4900001	5010000	
WP	ACF67367_50	5000001	5110000	
WP	ACF67367_51	5100001	5210000	
WP	ACF67367_52	5200001	5310000	
WP	ACF67367_53	5300001	5410000	
WP	ACF67367_54	5400001	5510000	
WP	ACF67367_55	5500001	5610000	
WP	ACF67367_56	5600001	5648894	
Query Match 10.0%; Score 81.6; DB 7; Length 110000;				
Best Local Similarity 52.6%; Pred. No. 1.1e-11;				
Matches 202; Conservative 0; Mismatches 179; Indels 3; Gaps 1;				
QY	256	ACAAAGTGGAGCAATTGTCAGACCAATCGGAAATCCAGCAAGTTTCATCGGCTTTACTGA	315	
DB	49307	AAAAGATTGATCGGATTATCAAACTTTCAAAATTAGATGATGTGGTGAAGCTCTGGCGG	49366	
QY	316	AAATCGGATTGAGGTGTTACTGTTCTGATGTGAGAGGGTTTGGTGCAAGGAGGTT	375	
DB	49367	AAGTGGGTATCACCGGAATGACAGTAACAGAGGTGAAAGGTTTGGGCGCCAAAAGGTC	49426	
QY	376	CTACCGAGAGACACGGTGGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGA	435	
DB	49427	ATACAGAGCTGTATCGCGGTGCAGAAATATATGTTGGAT---TTTCTGCCAAAAGTGAAA	49483	
QY	436	TGGAATCTGTTTAAAGAAAGACCAAGTGAATCTGTAATCAACACAAATATTAAGAGGAG	495	
DB	49484	TAGAAATTGCTGCGGACAGATGATATTGTCGATACCTGTGTGAAACCAATATGCAGACGG	49543	
QY	496	CAAGGACAGAGAGATTGGTGATGGCAAGATTTTGGTCCGTCGATGTCATATAA	555	
DB	49544	CACAGACCGGAAAATCGGTGATGGTAAATATTTGATTTGATGTAGCACGTTGTGTC	49603	
QY	556	GAGTTAGGACAGAGATTGGTGATGGCAAGATTTTGGTCCGTCGATGTCATATAA	615	
DB	49604	GTATCCGCACCGGTGAGCAGGATGAGAGGCGGATTAACCTTAATTTGAATTTTGGGGT	49663	
QY	616	CGTCTTAGGAACAAACAGAGCTCA	639	
DB	49664	GGTGACAGGCAGATAAAAAAGCCA	49687	
RESULT 11				
ACF65387_3				
Continuation (4 of 7) of ACF65387 from base 300001 (Photorhabdus luminescens nucleotide				
WP Sequence split into 7 fragments LOCUS ACF65387 Accession ACF65387				
WP	Fragment Name	Begin	End	
WP	ACF65387_0	1	110000	
WP	ACF65387_1	100001	210000	
WP	ACF65387_2	200001	310000	
WP	ACF65387_3	300001	410000	
WP	ACF65387_4	400001	510000	

WP	ACF65387_5	500001	610000	
WP	ACF65387_6	600001	696798	
Query Match 10.0%; Score 81.6; DB 7; Length 110000;				
Best Local Similarity 52.6%; Pred. No. 1.1e-11;				
Matches 202; Conservative 0; Mismatches 179; Indels 3; Gaps 1;				
QY	256	ACAAAGTGGAGCAATTGTCAGACCAATCGGAAATCCAGCAAGTTTCATCGGCTTTACTGA	315	
DB	38821	AAAAGATTGATCGGATTATCAAACTTTCAAAATTAGATGATGTGGTGAAGCTCTGGCGG	88880	
QY	316	AAATCGGATTGAGGTGTTACTGTTCTGATGTGAGAGGGTTTGGTGCAAGGAGGTT	375	
DB	88881	AAGTGGGTATCACCGGAATGACAGTAACAGAGGTGAAAGGTTTGGGCGCCAAAAGGTC	88940	
QY	376	CTACCGAGAGACACGGTGGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGA	435	
DB	88941	ATACAGAGCTGTATCGCGGTGCAGAAATATATGTTGGAT---TTTCTGCCAAAAGTGAAA	88997	
QY	436	TGGAATCTGTTTAAAGAAAGACCAAGTGAATCTGTAATCAACACAAATATTAAGAGGAG	495	
DB	88998	TAGAAATTGCTGCGTCCAGATGATATTGTCGATACCTGTGTGAAACCAATATGCAGACGG	89057	
QY	496	CAAGGACAGAGAGATTGGTGATGGCAAGATTTTGGTCCCTGTCAGATGTCATAA	555	
DB	89058	CACAGACCGGAAAATCGGTGATGGTAAATATTTGATTTGATGTAGCACGTTGTGTC	89117	
QY	556	GAGTTAGGACAGGTGAGCGTGGGAGAAAGCAGAGAAGATGACTGGTGATATGCTTTTAC	615	
DB	89118	GTATCCGCACCGGTGAGCAGGATGAGAGGCGGATTAACCTTAATTTGAATTTTGGGGT	89177	
QY	616	CGTCTTAGGAACAAACAGAGCTCA	639	
DB	89178	GGTGACAGGCAGATAAAAAAGCCA	89201	

RESULT 12

ACF72027	ID	ACF72027	standard; DNA; 339 BP.
XX	AC	ACF72027;	
XX	DT	20-NOV-2003	(first entry)
XX	DE	Photorhabdus luminescens	nucleotide sequence #10494.
XX	KW	Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;	
XX	KW	detection; food; gene expression; plant; animal; microorganism; toxin;	
XX	KW	antibiotic; biopesticide; virulence factor; disease model; plague;	
XX	OS	whooping cough; gene; ds.	
XX	PN	Photorhabdus luminescens.	
XX	PD	WO200294867-A2.	
XX	PF	28-NOV-2002.	
XX	PR	07-FEB-2002; 2002WO-IB003040.	
XX	PA	07-FEB-2001; 2001FR-00001659.	
XX	PA	(INSP) INST PASTEUR.	
XX	PA	(CNRS) CNRS CENT NAT RECH SCI.	
PI	PI	Duchaud B, Taourit S, Glaser P, Frangeul L, Kunst F, Darchin A;	
XX	PI	Buchrieser C;	
XX	DR	WPI; 2003-148459/14.	
XX	PT	Genomic sequence of Photorhabdus luminescens and encoded polypeptides,	
XX	PT	useful e.g. as therapeutic antimicrobials and agricultural pesticides.	
PS	PS	Claim 2; SEQ ID NO 10494; 1205pp; French.	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 19:07:55 ; Search time 1892.7 Seconds
(without alignments)
12890.278 Million cell updates/sec

Title: US-09-756-541-13
Perfect score: 817
Sequence: 1 CTGAAAGTTGTGTTAAAAA.....GTTCTAATAAAAAAAA 817

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hcc.*
9: gb_est1.*
10: gb_est2.*
11: gb_hcc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	499.6	61.2	566	9 AU236084	AU236084 AU236084
2	498.4	61.0	549	14 CB261070	CB261070 33-B9570-
3	459.6	56.3	685	14 CD825321	CD825321 BN25.060G
4	458.8	56.2	732	14 CD820903	CD820903 BN25.040D

5	456.8	55.9	745	14 CD819046	CD819046 BN20.047L
6	454	55.6	669	14 CD822497	CD822497 BN25.045G
C 7	369.6	45.2	451	9 AU226889	AU226889 AU226889
8	270.6	33.1	782	14 CB289081	CB289081 V-B-112F0
C 9	262.8	32.2	723	13 BQ798001	BQ798001 EST 6939
10	252.8	30.9	732	9 AJ558383	AJ558383 AJ558383
C 11	249.6	30.6	812	14 CA922829	CA922829 EST640547
12	247.8	30.3	714	12 BI311072	BI311072 EST531282
13	246.6	30.2	597	10 BE239550	BE239550 EST403599
14	245	30.0	613	9 AW035791	AW035791 EST281945
15	245	30.0	702	12 BI932123	BI932123 EST552012
16	245	30.0	711	12 BI932336	BI932336 EST552225
17	245	30.0	736	12 BI932913	BI932913 EST552802
18	245	30.0	741	10 AW738071	AW738071 EST339498
19	245	30.0	804	12 BI930336	BI930336 EST550225
20	245	30.0	810	12 BI930060	BI930060 EST549949
21	244.8	30.0	612	10 BF645522	BF645522 NF036F08E
22	244	29.9	648	9 AI773079	AI773079 EST254179
23	243.4	29.8	630	10 AW929534	AW929534 EST338322
24	243.4	29.8	685	12 BI929601	BI929601 EST549490
25	241.8	29.6	799	12 BM113473	BM113473 EST561009
C 26	240.8	29.5	677	10 BE823431	BE823431 GM700019B
27	240.2	29.4	756	12 BG595705	BG595705 EST494383
28	240.2	29.4	759	14 CF920467	CF920467 gmrhRwW3-
29	239.4	29.3	458	10 BE210153	BE210153 so39d03.Y
30	231.8	28.4	560	14 CA514765	CA514765 KS09030G0
31	231.8	28.4	579	10 AW223643	AW223643 EST300454
C 32	224.8	27.5	675	14 CB346213	CB346213 CAB2SG000
33	224.6	27.5	866	13 C81819	C81819 C81819 Citr
C 34	223.2	27.3	719	12 BJ573686	BJ573686 BJ573686
35	222.4	27.2	540	14 CB349912	CB349912 CAB2SG000
C 36	221.8	27.1	551	14 CB349990	CB349990 CAB2SG000
C 37	221.8	27.1	598	14 CB350294	CB350294 CAB2SG000
38	221.6	27.1	657	14 CB349285	CB349285 CAB2SG000
C 39	221.6	27.1	701	14 CB349370	CB349370 CAB2SG000
40	219.2	26.8	540	10 AW624562	AW624562 EST322507
41	218.6	26.8	532	14 CB350222	CB350222 CAB2SG000
42	213	26.1	564	13 BQ118918	BQ118918 EST604494
C 43	213	26.1	729	14 CF475242	CF475242 RTWW2_14
44	212.8	26.0	796	14 CF395143	CF395143 RTDS2_9_G
45	211.8	25.9	673	14 CB085731	CB085731 hg15b09.g

ALIGNMENTS

RESULT 1
AU236084
LOCUS AU236084 566 bp mRNA linear EST 01-APR-2002
DEFINITION AU236084 RAFL14 Arabidopsis thaliana cDNA clone RAFL14-63-K15 5', mRNA
mRNA sequence.
ACCESSION AU236084
VERSION AU236084.1 GI:19875253
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 566)
AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.
TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL Unpublished (2002)
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp

An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda ELC-1 vector (Carninci et al., submitted for publication) digested with BamHI and Sali. This clone is in a modified pBluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES		Location/Qualifiers	
source		1..566	
		/organism="Arabidopsis thaliana"	
		/mol_type="mRNA"	
		/db_xref="taxon:3702"	
		/clone="RAFL14-63-K15"	
		/tissue_type="root"	
		/lab_host="DH10B"	
		/clone_lib="RAFL14"	
		/note="Site_1: BamHI; Site_2: Sali"	
ORIGIN			
Query Match		61.2%; Score 499.6; DB 9; Length 566;	
Best Local Similarity		98.2%; Pred. No. 4.8e-94;	
Matches 536; Conservative		0; Mismatches 6; Indels 4; Gaps 3;	
QY	1	CTGAAAGTTGTGTTAAATAAACTAGATCATGGC-GCGGTCAATGACGAAACCCATC	59
Db	20	CTGAAAGTTGTGTTAAATAAACTAGATCATGGCAGCGTCAATGACGAAACCCATC	79
QY	60	TCAATAACTTCTCTCGGTTTCTATTCTGATCGAAAGAACATTGCTTCTCTGATTGCATT	119
Db	80	TCAATAACTTCTCTCGGTTTCTATTCTGATCGAAAGAACATTGCTTCTCTGATTGCATT	139
QY	120	TCGATTGTTCTGGATTTCAGACATTCCTGATCGAAAGAACATTGCTTCTCTGATTGCATT	179
Db	140	TCGATTGTTCTGGATTTCAGACATTCCTGATCGAAAGAACATTGCTTCTCTGATTGCATT	199
QY	180	CCGAGTAAATACAGTCGTGTTTACCTGTCGTAGTGCCCAATATCTTCTGATTATATT	239
Db	200	CCGAGTAAATACAGTCGTGTTTACCTGTCGTAGTGCCCAATATCTTCTGATTATATT	259
QY	240	CCAGACTCGAAATTTTACAAGGTGGAAGCAATTTGTGACACCATGAGAAATCCAGCAAGTT	299
Db	260	CCAGACTCGAAATTTTACAAGGTGGAAGCAATTTGTGACACCATGAGAAATCCAGCAAGTT	319
QY	300	TCATCGGCTTTACTGAAATCGGGATTCGAGGTGTTACTGTTTCTGATGTGAGAGGGTTT	359
Db	320	TCATCGGCTTTACTGAAATCGGGATTCGAGGTGTTACTGTTTCTGATGTGAGAGGGTTT	379
QY	360	GGTGCAACAGGAGGTTCTACCGAGAGACACGGTGGCTCTGAGTTCTCGAAGACAAATTT	419
Db	380	GGTGCAACAGGAGGTTCTACCGAGAGACACGGTGGCTCTGAGTTCTCGAAGACAAATTT	439
QY	420	GTTGCTAAAGTTAAGATGGAATCGTTGTTAAGAAAGACCAAGT-GGAATCTGTAATCAA	478
Db	440	GTTGCTAAAGTTAAGATGGAATCGTTGTTAAGAAAGACCAAGTGGGAATCTGTAATCAA	499
QY	479	CACAATAATTGAAGGAGCAAGG--ACAGGAGAGATTGGTGATGGCAAGATTTTGTGTTTG	536
Db	500	CACAATAATTGAAGGAGCCANGGACAGGAGAGATTGGTGATGGCAAGATTTTGTGTTTG	559
QY	537	CCTGTG 542	
Db	560	CCTGTG 565	
RESULT 2			
LOCUS CB261070			
DEFINITION 33-E9570-012-004-A09-T7R MP1Z-ADIS-012 Arabidopsis thaliana cDNA			
clone MP1Zp769A094Q 5-PRIME, mRNA sequence.			
ACCESSION CB261070			
VERSION CB261070.1 GI:32885843			
KEYWORDS EST.			
SOURCE Arabidopsis thaliana (thale cress)			

ORGANISM	Arabidopsis thaliana	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
REFERENCE	1 (bases 1 to 549)	
	Schmid,K.J., Soerensen,T.R., Stracke,R., Torjek,O., Altmann,T., Mitchell-Olds,T. and Weissshaar,B.	
AUTHORS	Large-scale identification and analysis of genome-wide single-nucleotide polymorphisms for mapping in Arabidopsis thaliana	
	Genome Res. 13 (6), 1250-1257 (2003)	
TITLE	ADIS DNA core facility at MPIZ	
	Max-Planck-Institute for Plant Breeding Research	
JOURNAL	Carl-von-Linne Weg 10, 50829 Koeln, Germany	
	Fax: 00492215062851	
MEDLINE	Email: weisshaar@mpiz-koeln.mpg.de	
	Insert length: 549 Std Error: 0.00	
PUBMED	Plate: 4 row: A column: 09	
	Seq primer: T7R; CTAATACGACTCACTATAGGGA.	
COMMENT	Location/Qualifiers	
	1..549	
FEATURES	/organism="Arabidopsis thaliana"	
	/mol_type="mRNA"	
source	/cultivar="Landsberg erecta (Ler)"	
	/db_xref="GABI:589834"	
QY	/db_xref="taxon:3702"	
	/clone="MP1Zp769A094Q"	
QY	/tissue_type="whole plant"	
	/dev_stage="adult plant, mixed stresses"	
QY	/lab_host="E. coli XL1-Blue MRF"	
	/clone_lib="MPIZ-ADIS-012"	
QY	/note="Vector: pSPORT1; Site 1: Sali; Site 2: NotI; cDNA library from Arabidopsis thaliana, accession Landsberg erecta; six weeks old total plants grown under long-day conditions in soil, whole adult plants were treated for 24 hours with different stresses, (1) at 4M-0 C in the dark, (2), at 37 Grad C in the dark, (3) lying in the lab after removing from soil, (4) in the greenhouse after wounding leaves with a foreceps, (5) in the lab watering with a 150 mM NaCl solution, (6) at 26 Grad C in the light/UV; equal quantities of stressed plant material were pooled; library was made at the Max-Planck-Institute for Plant Breeding Research, Cologne, Germany; cloning sites Sali-NotI, primer sites and orientation:	
	T7-Sali-CCACGGTCCG-5prime-cDNA-polyA-CC-NotI-SP6; Note: Sequencing granted in the context of the GABI Arabidopsis Verbund I: Genetic Diversity, 'Establishment of high-efficiency SNP-based mapping tools and development of methods for genome-wide mutation detection' PI: Bernd Weissshaar Sequence submission managed by RZPD/GABI-Primary database: http://gabi.rzpd.de . This clone is available from RZPD; contact RZPD (clone@rzpd.de) for further information."	
ORIGIN	Query Match	
	Best Local Similarity 97.6%; Pred. No. 8.6e-94;	
QY	14	TAAAAAACTAGATCATGGCGGCTCAATGACGAAACCCATCTCAATAACTTCTCT 73
	11	TAAAAAACTAGATCATGGCGGCTCAATGACGAAACCCATCTCAATAACTTCTCT 70
QY	74	CGGTTTCTATTCTGATCGAAAGAACATTGCTTCTCTGATTGCAATTCGATTTGTTCTGG 133
	71	CGGTTTCTATTCTGATCGAAAG-AACATTGCTTCTCTGATTGCAATTCGATTTGTTCTGG 129
QY	134	ATTCAGACATTCCCGACCATCTTGCCCTCGATTGTCACAAAGTCACCGAGTAATAACAG 193
	130	ATTCAGACATTCTCGACCATCTTGCCCTCGATTGTCACAAAGTCACCGAGTAATAACAC 189

QY 194 TCGTGTTTTACCTGTCGTTAGTCCCAATAATCTTCTGATTATATTCAGACTC-GAAAT 252
|||||
Db 190 TCGTGTTTTACCTGTCGTTAGAGCCCAAGCTCTTCTGATTATATTCAGACTCAGAAAT 249
|||||
QY 253 TTTACAAGGTGGAAGCAATTGTCAGACCATGGAGAATCCAGCAAGTTTCATCGGCTTTAC 312
|||||
Db 250 TTTACAAGGTGGAAGCAATTGTCAGACCATGGAGAATCCAGCAAGTTTCATCGGCTTTAC 309
|||||
QY 313 TGAATAATCGGGATTTCGAGGTGTTACTGTTTCTGATGTGAGAGGGTTGGTGCAAGGAG 372
Db 310 TGAATAATCGGGATTTCGAGGTGTTACTGTTTCTGATGTTCGAGGGTTGGTGCAAGGAG 369
|||||
QY 373 GTTCTACCGAGAGACACGGTGGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTA 432
Db 370 GTTCTACTGAGAGACACGGTGGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTA 429
|||||
QY 433 AGATGGAAATCGTTGTTAAGAAAGACCAAGTGGAAATCTGTAATCAACACAATTAATTGAAG 492
Db 430 AGATGGAAATCGTTGTTAAGAAAGACCAAGTGGAAATCTGTAATCAACACAATTAATTGATG 489
|||||
QY 493 GAGCAAGCAGCAGGAGAGATTGGTGATGGCAAGATTTTGTGTTTGCCTGTGTGATGTCA 552
Db 490 GAGCAAGCAGCAGGAGAGATTGGTGATGGCAAGATTTTGTGTTTGCCTGTGTGATGTCA 549
|||||

RESULT 3
CD825321
LOCUS
DEFINITION
sequence.
CD825321
CD825321.1 GI:32507261
EST.
Brassica napus (rape)
Brassica napus
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 685)
Genoplante.
Genoplante, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (<http://www.genoplante.com>
and <http://genoplante-info.infobiogen.fr>).
Location/Qualifiers
1 .685
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Jet neuf"
/db_xref="taxon:3708"
/clone="BN25060G17"
/tissue_type="seed"
/clone_lib="BN25"

FEATURES
source
1 .685
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Jet neuf"
/db_xref="taxon:3708"
/clone="BN25060G17"
/tissue_type="seed"
/clone_lib="BN25"

ORIGIN
Query Match 56.3%; Score 459.6; DB 14; Length 685;
Best Local Similarity 85.9%; Pred. No. 9.5e-86;
Matches 534; Conservative 0; Mismatches 84; Indels 4; Gaps 2;

QY 29 AATCATGGCGGCTCAATGACGAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGA 88
|||||
Db 5 AATCATGGCGGCTTCAATCGCGGAACCCATCTCTGCAACCTCTCTCGGTTTCTACTCCGA 64
|||||
QY 89 TCGAAAG---AACATGCTTTTCTCTGATTGCAATTCGATTTGTTCTGGATTACAGCATTC 145
|||||
Db 65 CCGAAGAACACAACATTTCCCTTCTCAGATTGCAATCTCGATATCTTCTGGTTCGACATTC 124
|||||

QY 146 CCGACCACTCTGCTCGATTTGGTCACAAAGTCACCGAGTAATAACAGTCGTGTTTTACC 205
|||||
Db 125 CCCACCGTCTTCCCTCGATTTGGTCGCAAAACCCACCACCGACCAAGGCTCGCGTTCCCA 184
|||||
QY 206 TGTCGTTAGTCCCAAAATATCTTCTGATTATATTCAGACTCGAAATTTTACAAGGTGGA 265
|||||
Db 185 TCTCGTTAGAGCCCAAGCTCTTATGATTATATTCAGACTCAAAGTTTACAAGGTGGA 244
|||||
QY 266 AGCAATTGTCAGACCATGGAGAATCCAGCAAGTTTCATCGGCTTTACTGAAAATCGGGAT 325
Db 245 AGCAATTGTCAGGCCATGGAGAATCCAGCATGTTTCATCGGCTTTACTGAAAATTTGGAT 304
|||||
QY 326 TCGAGGTGTTACTGTTTCTGATGTGAGAGGGTTGGTGCAAGAGGTTCTACCGAGAG 385
Db 305 TCGAGGTGTGACGGTTTCTGATGTTCAGAGGGTTGGTGCAAGGTGGATCCAAAGAGAG 364
|||||
QY 386 ACACGGTGGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGGAAATCGT 445
Db 365 ACATGGTGGTCTGAGTTCTCTGAAGACAAAGTTGTTGCTTAAGTTAAGATGGAGATCGT 424
|||||
QY 446 TGTTAAGAAAGACCAAGTGGAAATCTGTAATCAACACAATAATTGAAGGAGCAAGGACAGG 505
Db 425 TGTTAAGAAAGACCAAGTGGAGTCTGTAATCAACACCATAATCGATGGAGCAAGAACAGG 484
|||||
QY 506 AGAGATTGCTGATGCGCAAGATTTTGTGTTTGCCTGTGTGATGTCTATAAGAGTTAGGAC 565
Db 485 AGAGATCGGAGATGGCAAAATATTGCTGTTTGCCTGTGTGATGTCTATAAGAGTTAGGAC 544
|||||
QY 566 AGGTGAGCGTGGGAGAAAGCAGAGAGATGACTGGTGATATGCTTTCACCGCTCTTAGGA 625
Db 545 AGGTGAGCGTGGGAGAAACAGCAGAGAAGATGACTGGTGACATGCTTTCGTCACTTAGGC 604
|||||
QY 626 ACAAAAC-AGAGCTCAAGAANTGG 646
Db 605 ACAAAACAGAGCTTAAGAANTGG 626
|||||

RESULT 4
CD820903
LOCUS
DEFINITION
sequence.
CD820903
CD820903.1 GI:32502843
EST.
Brassica napus (rape)
Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 732)
Genoplante.
Genoplante, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (<http://www.genoplante.com>
and <http://genoplante-info.infobiogen.fr>).
Location/Qualifiers
1 .732
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Jet neuf"
/db_xref="taxon:3708"
/clone="BN25040D05"
/tissue_type="seed"
/clone_lib="BN25"

FEATURES
source
1 .732
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Jet neuf"
/db_xref="taxon:3708"
/clone="BN25040D05"
/tissue_type="seed"
/clone_lib="BN25"

ORIGIN
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (<http://www.genoplante.com>
and <http://genoplante-info.infobiogen.fr>).
Location/Qualifiers
1 .732
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Jet neuf"
/db_xref="taxon:3708"
/clone="BN25040D05"
/tissue_type="seed"
/clone_lib="BN25"


```
RESULT 6
CD822497
LOCUS
DEFINITION
  CD822497 BN25.045G20F0201c8 BN25 Brassica napus cDNA clone BN25045G20, mRNA
  sequence.
ACCESSION
CD822497
VERSION
CD822497.1 GI:32504437
KEYWORDS
EST.
SOURCE
  Brassica napus (rape)
  Brassica napus
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
  1 (bases 1 to 669)
  Genoplatte.
  Genoplatte, a major partnership french program in plant genomics
  Unpublished (2003)
  Contact: Genoplatte
  Genoplatte
  93, rue Henri Rochefort 91025 EVRY CEDEX France
  Tel: 33 1 69 47 54 00
  Fax: 33 1 69 47 54 10
  This sequence has been generated in the framework of the french
  plant genomics programme 'Genoplatte' (http://www.genoplatte.com
  and http://genoplatte-info.infobiogen.fr).
  Location/Qualifiers
    1..669
      /organism="Brassica napus"
      /mol_type="mRNA"
      /cultivar="Jet neuf"
      /db_xref="taxon:3708"
      /clone="BN25045G20"
      /tissue_type="seed"
      /clone_lib="BN25"

ORIGIN
Query Match 55.6%; Score 454; DB 14; Length 669;
Best Local Similarity 85.6%; Pred. No. 1.4e-84;
Matches 529; Conservative 0; Mismatches 85; Indels 4; Gaps 2;

33 ATGGCGGCGTCAATGACGAACCCATCTCAATAACTTCTCGGTTTCTATTCGATCGA 92
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2 ATGGCGGCTTCAATCGGAACCCATCTCTGCAACCTCTCTCCGTTTCTACTCCGACCGA 61
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
93 AAG---AACATTGCTTCTCTGATTGCAATTCGATTGTTCTGGATTTCAGACATCCCGA 149
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
62 AAGAACCAACATTTCTCTCAGATTGCATCTCGATATCTCTGGTTCCGACATTCGCCA 121
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
150 CCATCTTGCTCGATTTCGTCACAAAGTCACCGAGTAATAACAGTCGTGTTTACCTGTC 209
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
122 CGCTCTCCCTCGATTGTCGCAAAACCAACCCACCGACCAAGCTCGCGTTTCCCATCTC 181
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
210 GTTAGTGCCCAAAATCTTCTGATTATATTCAGACTCGAAATTTTACAAGGTGGAAGCA 269
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
182 GTTAGAGCCCAAGCTCTTATGATTATATTCAGACTCAAAGTTTACAAGGTGGAAGCA 241
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
270 ATTGTCAGACCATGGAATCCAGCAAGTTTTCATCGSCTTTACTGAAATCGGATTGGA 329
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
242 ATTGTCAGGCCATGGAATCCAGCATGTTTCATCGGCTTTACTGAAATTCGGATTGGA 301
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
330 GGTGTTACTGTTTCTGATGTGAGAGGGTTTGTCGCAAGGAGGTTCTACCGAGAGACAC 389
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
302 GGTGTGACGGTTTCTGATGTCCGAGGGTTTGTGTGCAAGGTGGATCCAAAGAGAGACAT 361
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
390 GGTGGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTT 449
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
362 GGTGGTCTGAGTTCTCTGAAGACAAGTTCGTGCTAAAGTTAAGATGGAATCGTTGTT 421
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
450 AAGAAAGACCAAGTGGAAATCTGTAATCAACACAATAATTGAAGGAGCAAGGAGAGAG 509
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
422 AAGAAAGAACAAAGTGGAGTCTGTAATCAACACCATAATCGATGGAGCAAGAACAGGAG 481
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
510 ATTGGTATGGCAAGATTTTGTGTTTTCCTGTGTGATGTCTAATAGAGTTAGGACAGGT 569
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

FEATURES
source
  Location/Qualifiers
    1..451
      /organism="Arabidopsis thaliana"
      /mol_type="mRNA"
      /db_xref="taxon:3702"
      /clone="RAFL14-63-K15"
      /tissue_type="root"
      /lab_host="DH10B"
      /clone_lib="RAFL14"
      /note="Site_1: BamHI; Site_2: SalI"

ORIGIN
Query Match 45.2%; Score 369.6; DB 9; Length 451;
Best Local Similarity 96.9%; Pred. No. 6.3e-67;
Matches 409; Conservative 0; Mismatches 9; Indels 4; Gaps 3;

391 GTGGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTA 450
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
450 GTGGCTCTGAGTTCTC-GAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTA 392
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
451 AGAAAGACCAAGTGGAAATCTGTAATCAACACAATAATTGAAGGAGCAAGGAGAGAGA 510
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
391 AGAAAGACCAACTGGAAATCTGTAATCAACACAATAATTCAAGGAGC-ACGACAGGAGAG 333
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
511 TTGGTGTATGGCAAGATTTTGTGTTTTCCTGTGTGATGTCTAATAGAGTTAGGACAGGTG 570
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
332 TTGGTGTATGGCAAAATTTTGTGTTTTCCTGTGTGATGTCTAATAGAGTTAGGACAGGTG 273
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db
482 ATCGGAGATGGCAAAATATTCGTTTTCCTGCTGTCAGATGTCATAAGGGTTCGGACAGGT 541
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY
570 GAGCGTGGGAGAAAGCAGACAGACATGACTGGTGATATGCTTTCACCGTCTTAGGAACAA 629
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db
542 GAGCGTGGAGAACACAGCAGAGAGATGACTGGTGACATGCTTTCGTCATCTTAGGCACAA 601
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY
630 AC-AGAGCTCAAGATGG 646
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db
602 ACAAGAGCTTAAGATGG 619
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
AU226889/c
LOCUS
DEFINITION
  AU226889 RAFL14 Arabidopsis thaliana cDNA clone RAFL14-63-K15 3',
  mRNA sequence.
ACCESSION
AU226889
VERSION
AU226889.1 GI:19741536
KEYWORDS
EST.
SOURCE
  Arabidopsis thaliana (thale cress)
  Arabidopsis thaliana
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
  1 (bases 1 to 451)
  Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
  Akiyama,K., Enju,A., Oono,Y., Sakurai,T., Carninci,P., Kawai,J.,
  Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A.,
  Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.
  Large scale analysis of Arabidopsis full-length cDNA
  Unpublished (2002)
  Contact: Motoaki Seki
  Plant Functional Genomics Research Group
  RIKEN Genomic Sciences Center
  3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
  Tel: 81-298-36-4359
  Fax: 81-298-36-9060
  Email: mseki@rtc.riken.go.jp
  An Arabidopsis full-length cDNA library was constructed essentially
  as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
  and XhoI was ligated to modified lambda FLC-1 vector (Carninci et
  al., submitted for publication) digested with BamHI and SalI. This
  clone is in a modified pBluescript vector. Please visit our web
  site (http://www.gsc.riken.go.jp/e/plant/index\_e.html) for further
  details.
  Location/Qualifiers
    1..451
      /organism="Arabidopsis thaliana"
      /mol_type="mRNA"
      /db_xref="taxon:3702"
      /clone="RAFL14-63-K15"
      /tissue_type="root"
      /lab_host="DH10B"
      /clone_lib="RAFL14"
      /note="Site_1: BamHI; Site_2: SalI"

ORIGIN
Query Match 45.2%; Score 369.6; DB 9; Length 451;
Best Local Similarity 96.9%; Pred. No. 6.3e-67;
Matches 409; Conservative 0; Mismatches 9; Indels 4; Gaps 3;

391 GTGGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTA 450
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
450 GTGGCTCTGAGTTCTC-GAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTA 392
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
451 AGAAAGACCAAGTGGAAATCTGTAATCAACACAATAATTGAAGGAGCAAGGAGAGAGA 510
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
391 AGAAAGACCAACTGGAAATCTGTAATCAACACAATAATTCAAGGAGC-ACGACAGGAGAG 333
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
511 TTGGTGTATGGCAAGATTTTGTGTTTTCCTGTGTGATGTCTAATAGAGTTAGGACAGGTG 570
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
332 TTGGTGTATGGCAAAATTTTGTGTTTTCCTGTGTGATGTCTAATAGAGTTAGGACAGGTG 273
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

QY 571 AG--CGTGGGAGAAAGCAGAGAGATGACTGGTGATATGCTTTCACCGTCTTAGGAACA 628
Db |||||
272 AGGCGTGGGGAGAAAGCAGAGAGATGACTGGTGATATGCTTTCACCGTCTTAGGAACA 213
QY 629 AACAGAGCTCAAGAATGGTTTTTTTTCATTTTCGCTCTAGATTCTGCGAATAAT 688
Db |||||
212 AACAGAGCTCAAGAATGGTTTTTTTTCATTTTCGCTCTAGATTCTGCGAATAAT 153
QY 689 AATGAATGGAGTCTGTGTTTGGTTTCATGTTGAATCGATCAAGATGTGTTTAACTGTA 748
Db |||||
152 AATGAATGGAGTCTGTGTTTGGTTTCATGTTGAATCGATCAAGATGTGTTTAACTGTA 93
QY 749 CATGAATATGAGAAACATCTGCTCTGTTCTCAGACATCGAAACTCTGTTCTTAATAA 808
Db |||||
92 CATGAATATCCAGAAACATCTGCTCTGTTCTCAGACATCGAAACTCTGTTCTTAATAA 33
QY 809 AA 810
Db ||
32 AA 31

RESULT 8
CB289081
LOCUS
DEFINITION V-B-112F06 VAN-Baker-1 Vitis aestivalis cDNA clone V-B-112F06 5',
mRNA sequence.
ACCESSION CB289081
VERSION CB289081.1 GI:28602822
KEYWORDS EST.
SOURCE Vitis aestivalis
ORGANISM Vitis aestivalis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.
1 (bases 1 to 782)
Hou,H.S., Phanikanth,T.V., Kovacs,L. and Qiu,W.P.
Expressed sequence tags of young leaf tissues of a
disease-resistant Vitis aestivalis var. Norton
Unpublished (2003)
Contact: Wenping Qiu
Department of Fruit Science
Southwest Missouri State University-Mountain Grove
9740 Red Spring Road, Mountain Grove, MO 65711-2999, USA
Tel: 417 926 4105
Fax: 417 926 6646
Email: weq070f@smsu.edu
Insert Length: 782 Std Error: 0.00
Plate: VAN-Baker-1-12 row: F column: C6
Seq primer: T3 PRIMER
High quality sequence stop: 782
POLYA=No.

FEATURES
source
1..782
/organism="Vitis aestivalis"
/mol_type="mRNA"
/cultivar="Norton"
/db_xref="taxon:3605"
/clone="V-B-112F06"
/tissue_type="Leaf"
/dev_stage="Young leaf"
/lab_host="XL10-Gold E.coli"
/clone_lib="VAN-Baker-1"
/note="Vector: pBluescript II SK (+) Phagemid; Site 1: Xho
I; Site 2: EcoR I; VAN-Baker-1 is a cDNA library of Norton
grape young leaves (Vitis. aestivalis var. Norton). Norton
grapevines were grown under normal greenhouse conditions.
The cDNA synthesis and library construction was performed
according to the instruction manual for pBluescript II XR
cDNA library construction kit provided by Stratagene."

ORIGIN
Query Match 33.1%; Score 270.6; DB 14; Length 782;

Best Local Similarity 72.7%; Pred. No. 2.le-46;
Matches 364; Conservative 0; Mismatches 134; Indels 3; Gaps 1;
QY 103 CTTTCTCTGATTGCATTTCGATTTTGTCTGGATTTCAGACATTCCTCCGACCATCTTGCCTCG 162
Db |||||
74 CTCTCATAGACTGCATTTTGATCCGCCCAATCTCAAAGATTCTCGAAATTTCCAGTTTA 133
QY 163 ATTTGGTCAAAAAGTCACCGAGTAATAACAGTCGTGTTTTTACCTGTCGTAGTGCCTAAA 222
Db |||||
134 ATTTGTGCTTAAA--CCGCTCAAAGAATGGATCTGTTCTTCCCATGTTCAGAGCTCAGA 190
QY 223 TATCTTCTGATTATATTCAGACTCGAATTTTACAAGGTGGAAGCAATTCAGACCAT 282
Db |||||
191 GCTCTCCAGATTATACCCAGACGCGCAGTTTACAAAAGTTGAAGCGATCCTGAGGCCCT 250
QY 283 GGAGAAATCCAGCAAGTTTCATCGGCTTTACTGAAAATCGGATTCGAGGTGTTACTGTTT 342
Db |||||
251 GCGGAATCCAGCAGGTTTCTCGGCTTTGCTGAAAATGGGTATTCGTGGTGTACTGTTT 310
QY 343 CTGATGTGAGAGGGTTTGTGTCACAAGGAGGTTCTACCGAGAGACACGGTGGCTCTGAGT 402
Db |||||
311 CTGATGTCCGAGGCTTTGTGTCTCAAGCGGATCGCCTGAAAGACAGGCAGGCTCTGAAT 370
QY 403 TCTCGAAGACAAATTTTGTCTAAAGTTAAGTGAATCGTGTGTTTAAGAAAGACCAAG 462
Db |||||
371 TTTCTGAGGACAAATTTGTGCTAAAGTTAAATGGAGATTGTGTGAGCAAGACCAAG 430
QY 463 TGAATCTGTAATCAACAAATTAATGAAGGAGCAAGGAGAGATTGGTGTATGGCA 522
Db |||||
431 TTGAAGCAGTAATTGACAAGATCAATGAGGTGGCAAGGACTGGAGAGATTGGTGTATGGCA 490
QY 523 AGATTTTGTGTTTGCCTGTGTGATGATGTCATAGAGTTAGGACAGGTGAGCGTGGGAGA 582
Db |||||
491 AGATCTTTTGTGTCCTATATCGGATGTGTATAGAGTTTCGACCGGTGAACGTGGAGAGA 550
QY 583 AAGCAGAGAGAGATGACTGGTG 603
Db |||||
551 TGGCTGAGAGGATGACTGGTG 571

RESULT 9
BQ798001/c
LOCUS
DEFINITION BQ798001 723 bp mRNA linear EST 30-JUL-2002
vinifera cDNA clone RT093C02 3', mRNA sequence.
ACCESSION BQ798001
VERSION BQ798001.1 GI:22012967
KEYWORDS EST.
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.
1 (bases 1 to 723)
Abbal,P., Agasse,A., Ageorges,A., Atanassova,R., Barrieu,F.,
Couture,C., Dedaldechamp,F., Delrot,S., Glissant,D., Grimplet,J.,
Hamdi,S., Romieu,C. and Terrier,N.
Generation of Expressed Sequence Tag from Grape Berry (skin, pulp
or seeds) at Various Developmental Stages
Unpublished (2002)
Contact: Romieu C.
Unite de Recherche des Produits de la Vigne
Institut National de la Recherche Agronomique
2, place Viala, 34 060 Montpellier Cedex 01, France
Tel: 00-33-(0)4-99-61-28-62
Fax: 00-33-(0)4-99-61-28-57
Email: romieu@ensam.inra.fr
Seq primer: T7.

FEATURES
Location/Qualifiers
1..723
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Shiraz"

Baylor College of Medicine
1100 Bates Street, Houston, TX 77030-2600, USA
Tel: 713-798-7044
Fax: 713-798-7078
Email: mgrusak@bcm.tmc.edu
B397927e

TIGR sequence name: MTPAP50TK
More information is available at: www.medicago.org
seq primer: SKmod (CTA gAA CTA gtg gAT CC).

FEATURES
SOURCE

```

1. .714
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="A17"
/db_xref="taxon:3880"
/clone="pGESP9J4"
/tissue_type="immature seeds"
/dev_stage="immature seeds, 11 to 19 days after
pollination"

```

ORIGIN

```
Query Match      30.3%; Score 247.8; DB 12; Length 714;
Best Local Similarity 67.6%; Pred. No. 1.2e-41;
Matches 348; Conservative 0; Mismatches 167; Indels 0;
```

QY	199	TTTACCTGTCGTTAGTCCCAAAATATCTTCTGATTATATCCAGACTCGAAATTTTACA	250
Db	172	TTCTTCCCAAAATCAGAGCTCAAAAACCTTCCTGACTATGTTCTGAAATCCAAGTTTTCACA	231
QY	259	AGGTGGAAGCAATTTGTCAAGCATGGAGAAATCCAGCAAGTTTCATCGGCTTTACTGAAGA	318
Db	232	AAGTTGAAGCCATTCTCAGGCCATGGAGAAATCCCTCAGGTTTCTTCGGGTTTGTTCGAAGA	291
QY	319	TCGGGATTCGAGGTGTTACTGTTTCTGATGTGAGAGGGTTTGGTGCACAAGGAGGTTCTTA	378
Db	292	TGGGAATTCGTGGTGTCACTGTATCTGATGTCAAGGGGTTTGGTGTCTCAGGGTGGCTCAA	351
QY	379	CCGAGAGACACGGTGGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTTAAAGTTAAGATGG	438
Db	352	AAGAGAGGCAGGGAGGGCTCCGAAATTTTCTGAAGACAAATTTTGTGCCAAAGTTAAATGG	411
QY	439	AAATCGTTGTTAAGAAAGACCAAGTCGAATCTGTAATCAACACATTAATTGAAGGAGCAA	498
Db	412	AAATAGTGGTGAGAAAGACCAGGTTGAGGCAGTGATAACAACAATTTATCGAGACGGCAA	471
QY	499	GGACAGGAGAGATTGGTGATGGCAAGATTTTGTGTTTGGCTGTGTGAGTGTCTATAGAG	558
Db	472	GAACTGGGGAGATTGGTGATGGCAAAATTTTCTTGATCCCTGTATCTGATGTAAATAGAA	531
QY	559	TTAGGACAGGTGAGCGTGGGGAGAAAGCAGAGAGATGACTGGTGATATGCTTTCACCGT	618
Db	532	TCCGCACAGGTGAGCGTGGGGAGCAGGCTGAGAGGATGGCTGGGGACTAACTGACCGGT	591
QY	619	CTTAGGAACAAACAGAGCTCAAGATGTTTTTTTTTTTCAATTCGGTCTCTAGATTC	678
Db	592	TGTCGTGTTGACAAAAGATTAGCAATTTGCAATTTGGTCTCTGTTCTCTTTCCCTGTGA	651
QY	679	TGCGAATAAATGAATGGAGTCTGTGTTTGTTTTT	713

```

/db_xref="taxon:3880"
/clone="MTUS-58H7"
/tissue_type="mixed tissues"
/dev_stage="various stages"
/lab_host="XL0LR"
/clone_lib="MTUS"
/notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; CDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in XL0LR cells."

```

ORIGIN

Query Match	30.6%	Score 249.6	DB 14	Length 812
Best Local Similarity	70.6%	Pred. No. 5e-42		
Matches 333	Conservative	0	Mismatches 139	Indels 0
				Gaps 0

QY	199	TTTACCTGTCGTTAGTCCCAAAATATCTCTGATATATATCCAGACTCGAAATTTTACA	258
DB	623	TTCTTCCCAAAATCAGAGCTCAAAACCTTCCTGACTATGTTCTGAAATCCAAGTTTACA	564
QY	259	AGGTGGAAGCAATTGTCAGACCATGGAGAAATCCAGCAAGTTTTCATCGGCTTTACTGAAAA	318
DB	563	AAGTTGAAGCCAATCTCAGGCCATGGAGAAATCCCTCAGGTTTCTTCGGGTTTGTGAAAA	504
QY	319	TCCGGATTCGAGGTGTTACTGTTTCTGATGTGAGAGGGTTGGTGCAACAAGGAGGTTCTA	378
DB	503	TGGGAATTCGTGGTGTCACTGTATCTGATGTCAAGGGGTTGGTGCTCAGGGTGGCTCAA	444
QY	379	CCGAGAGACACCGTGGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGG	438
DE	443	AAGAGAGGCAGGGAGGCTCCGAATTTTCTGAAGACAAATTTGTTGCCAAAGTTAAATATGG	384
QY	439	AAATCGTTGTTAAGAAAGACCAAGTGGAACTCTGTAATCAACACAATAATTTGAAGGAGCAA	498
DB	383	AAATAGTGGTGAGAAAAGACCAAGGTTGAGGCAGTGATAAACAAAATTTATGAGACGGCAA	324
QY	499	GGACAGGAGAGATTGGTGATGGCAAGATTTTGTGTTTTCCTGTGTCTCAGATGTCATAAGAG	558
DB	323	GAACTGGGGAGATTGGTGATGGCAAAATTTTCTTGATCCCTGTATCTGATGTAATAAGAA	264
QY	559	TTAGGACAGGTGAGCGTGGGAGAAAGCAGAGAAGTGACTGGTGATATGCTTTTCACCGT	618
DB	263	TCCGCA CAGGTGAGCGTGGGGAGCAGGCTGAGAGGATGGCTGGGGGACTAACTGACGCGT	204
QY	619	CTTAGGAAACAAACAGAGCTCAAGAAATGGTTTTTTTTTTTTTTTTCATTTCGGTCT	670
DB	203	TGCTGTTTGAACAAAAGATTAGCAATTTGCAATTTGGTCCCTCTGTTCTCTTT	152

RESULT	12
BI311072	
LOCUS	
DEFINITION	BI3111072 714 bp mRNA linear EST 20-JUL-2001 EST5312822 GESD Medicago truncatula cDNA clone pGESD9J4 5' end, mRNA sequence.

ACCESSION BI311072
 VERSION BI311072.1 GI:14985399
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.

REFERENCE	1 (bases 1 to 714)
AUTHORS	Grusak, M.A., Samac, D.A., Town, C.D., Van Aken, S., Utterback, T., Cho, J., and Fraser, C.M.
TITLE	ESTs from developing reproductive tissues of <i>Medicago truncatula</i>
JOURNAL	Unpublished (2001)
COMMENT	Contact: Michael A. Grusak msa@aps.illinois.edu Nutrition Research Center


```
Db      652 GAATTGATAACAATGATGATGTTGGTGTGTTT 686

RESULT 13
LOCUS   BE239550
DEFINITION   MHRP- Medicago truncatula cDNA clone pMHRP-28F3, mRNA
sequence.
ACCESSION   BE239550
VERSION     BE239550.1 GI:9055733
KEYWORDS    Medicago truncatula (barrel medic)
SOURCE      Medicago truncatula
ORGANISM    Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1 (bases 1 to 597)
AUTHORS    Harrison,M.J., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S.,
            Holt,I.E., Cho,J. and Fraser,C.M.
TITLE      ESTs from phosphate-starved roots of Medicago truncatula
JOURNAL    Unpublished (2000)
COMMENT    Contact: Maria J. Harrison
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73401, USA
            Tel: 580-223-5810
            Fax: 580-221-7380
            Email: mjharrison@noble.org
            The Samuel Noble Roberts Foundation: N265417e
            TIGR sequence name:MTHAC26TK
            More information is available at.
            http://chrysie.tamu.edu/medicago
            Seq primer: SKmod (CTA GAA CTA gtg GAT CC).
            Location/Qualifiers
                1..597
                /organism="Medicago truncatula"
                /mol_type="mRNA"
                /cultivar="A17"
                /db_xref="taxon:3880"
                /clone="pMHRP-28F3"
                /tissue_type="roots"
                /dev_stage="phosphate-starved"
                /lab_host="XL0LR"
                /clone_lib="MHRP-"
                /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
                XhoI; At the trifoliolate stage, M. truncatula plants were
                transplanted to phosphate-free sand and grown for a
                further 30 days. During this period, they were fertilized
                twice weekly with 1/2 Hoaglands solutions containing 20uM
                potassium phosphate. cDNA was prepared from polyA+
                enriched RNA. The cDNA was directionally ligated into the
                Unizap XR vector from Stratagene and packaged using
                Gigapack III Gold packaging extracts. Plasmids containing
                cDNA inserts were excised from the recombinant lambda-Zap
                phage using Ex-assist helper phage and propagated in
                XL0LR cells."

ORIGIN
Query Match      30.2%; Score 246.6; DB 10; Length 597;
Best Local Similarity 75.6%; Pred. No. 2.4e-41;
Matches 306; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Qy      199 TTTTACCTGTCGTTAGTCCCAATATCTCTGATTATATTCAGACTCGAAATTTTACA 258
        |||||
Db      173 TTCTTCCCAAAATCAGAGCTCAAAACCTTCCTGACTATGTTCTGAATCCAAGTTTACA 232
        |||||

Qy      259 AGGTGAAGCAATTGTGAGACCATGAGAAATCCAGCAAGTTTCAFCGGCTTTACTGAAA 318
        |||||
Db      233 AAGTTGAAGCCATTCTCAGGCCATGAGAAATCCCTCAGGTTTCTTCGGTTTGTGAAA 292
        |||||

Qy      319 TCCGGATTGAGGTTGTTACTGTTCTGATGTGAGAGGGTTTGGTGCACAAGGAGTTCTA 378
        |||||
```

```
Db      293 TGGGAATTCGTGTCCTCACTGTATCTGATGTCAAGGGTTTGGTGCTCAGGGTGGCTCAA 352
Qy      379 CCGAGAGACACCGTGGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGG 438
        |||||
Db      353 AAGAGAGGCAGGAGGCTCCGAATTTTCTGAAGACAAATTTGTTGCCAAAGTTAAATGG 412
        |||||
Qy      439 AAATCGTTGTTAAGAAAGACCAGTGGAAATCTGTATATCAACACAATAATTGAAGGAGCAA 498
        |||||
Db      413 AAATAGTGTGTGAGAAAGACCAGGTTGAGCGCAGTGATAAACAAAATTATGGAGACCGCAA 472
        |||||
Qy      499 GGACAGGAGAGATTGTTGATGGCAAGATTTTTTTGTTGCTGTGTGATGTCTATAGAG 558
        |||||
Db      473 GAACTGGGGAGATTGTTGATGGCAAAATTTTCTTGATCCCTGTATCTGATGTAATAAGAA 532
        |||||
Qy      559 TTAGGACAGGTGAGCGTGGGAGAAAGCAGAGAGATGACTGGTG 603
        |||||
Db      533 TCCGCACAGGTGAGCGTGGGAGCAGGCTGAGAGGATGGCTGGGG 577

RESULT 14
LOCUS   AW035791
DEFINITION   EST281945 tomato callus, TAMU Lycopersicon esculentum cDNA clone
            cLEC36D8, mRNA sequence.
ACCESSION   AW035791
VERSION     AW035791.1 GI:5894547
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (tomato)
ORGANISM    Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
            1 (bases 1 to 613)
            Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
            Holt,I.E., Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S.,
            Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
            Giovannoni,J.
            Generation of ESTs from tomato callus tissue
            Unpublished (1999)
            Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
            Location/Qualifiers
                1..613
                /organism="Lycopersicon esculentum"
                /mol_type="mRNA"
                /cultivar="TA496"
                /db_xref="taxon:4081"
                /clone="cLEC36D8"
                /tissue_type="callus"
                /dev_stage="25-40 days old"
                /lab_host="XLI-Blue MRF"
                /clone_lib="tomato callus, TAMU"
                /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
                XhoI; supplier: Giovannoni laboratory; CLEC - Cotyledons
                of seedlings 7-10 days post-germination were excised, cut
                at both ends and placed on MS medium with no selection.
                Mixed callus was harvested at 25 and 40 days and included
                undifferentiated masses. Tomato Callus EST Library"

ORIGIN
Query Match      30.0%; Score 245; DB 9; Length 613;
Best Local Similarity 72.5%; Pred. No. 5.1e-41;
Matches 317; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

Qy      167 GGTCACAAAGTCACCGAGTAATAACAGTCGTGTTTACCTGTCGTTAGTCCCAATATC 226
        |||||
Db      131 GCTCACCTTCAACAGCTGTCAAATGCTCTCTTCCCGATTATCAGAGCCCAAACTC 190
        |||||

Qy      227 TTCTGATTATATCCAGACTCGAAATTTTACAAGGTGAAGCAATTTGTACACCATGGAG 286
        |||||
```

Db 191 TCCAGATTTCCTCCCTGATCGGAAGTTTACAAAGTTGAAGCAATCTAAGACCTTGGAG 250
QY 287 AATCCAGCAAGTTTCATCGCTTTACTGAAAATCGGATTCGAGGTGTTACTGTTCTGA 346
Db 251 AATCAACAGGTTTCTTCGGCACTACTGAAAATGGGCATTCGTGTCTCACTGTTCCGA 310
QY 347 TGTGAGAGGTTTGGTGCACAAGGAGTTCTACCGAGAGACACGCTGGCTCTGAGTTCTC 406
Db 311 TGTTCGTGTTTGGCGCCCAAGGTGGCTTGACTGAGAGGCAAGCTGGCTCTGAATCTC 370
QY 407 GGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAGAAAGACCAAGTGA 466
Db 371 TGAAGACACGTTTGTTCGCAAAAGTTAAATGGAATTTGTTGTCAGCAAGACCCAGGTTGA 430
QY 467 ATCTGTAATCAACACAATAATTGAAGGAGCAAGGACAGGAGAGATTGGTGATGGCAAGAT 526
Db 431 AGGAGTCATTGCCATGATAATTGAAGAGGCAAGAACTGGTGAAATAGGTGATGGAAGAT 490
QY 527 TTTTGTGTTGCTGTCAGATGTCATAAGAGTTAGGACAGCTGAGCGTGGGGAGAAAGC 586
Db 491 ATTCTTGACTCCCATCTCCGATGTTATAAGAGTTCCGCACTGCTGAACGGGGAGAAAGGC 550
QY 587 AGAGAGATGACTGGTG 603
Db 551 AGAGAGGATGATGGGAG 567

RESULT 15
BI932123
LOCUS
DEFINITION
EST552012 tomato flower, 8 mm to preanthesis buds Lycopersicon
esculentum cDNA clone cTOC22G23 5' end, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Lycopersicon esculentum (tomato)
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 702)
van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.

Generation of ESTs from tomato flower tissue, buds 8 mm -
preanthesis

Unpublished (2001)

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

This clone is available through the Clemson University Genomics

Institute

Seq primer: T3.

Location/Qualifiers

1..702
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOC22G23"
/tissue_type="flower"

/dev_stage="buds 8mm to preanthesis"
/clone_lib="tomato flower, 8 mm to preanthesis buds"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research; Flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

ORIGIN

Query Match 30.0%; Score 245; DB 12; Length 702;
Best Local Similarity 72.5%; Pred. No. 4.8e-41;
Matches 317; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
QY 167 GGTCAACAAAGTCACCGAGTAATAACAGTCGTGTTTACCTGTCGTTAGTCCCAATATC 226
Db 172 GCTCACCTTCAAACGCTGTCAAAATGCTCCTTCTTCCCGATTATCAGAGCCCAAACTC 231
QY 227 TTTGATTATATTCAGACTCGAAATTTTACAAGTGAAGCAATTTTCAGACCATGGAG 286
Db 232 TCCAGATTTGTCCTGATGCGAAGTTTACAAGTTGAAGCAATTTAAGACCTTGGAG 291
QY 287 AATCCAGCAAGTTTTCATCGGCTTTACTGAAAATCGGGATTCGAGGTGTTACTGTTCTGA 346
Db 292 AATTCAACAGGTTTCTTCGGCACTACTGAAAATGGGCAATTCGTGGTGTCACTGTTCCGA 351
QY 347 TGTGAGAGGTTTGGTGCCACAAGGAGGTTCTACCGAGAGACACGGTGGCTCTGAGTTCTC 406
Db 352 TGTTCGTGTTTGGCGCCCAAGGTGGCTTGACTGAGAGGCAAGCTGGTCTGAAATCTC 411
QY 407 GGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAGAAAGCAAGTGA 466
Db 412 TGAAGACACGTTTGTTCGCAAAAGTTAAATGGAATTTGTTGTCAGCAAGACCCAGGTTGA 471
QY 467 ATCTGTAATCAACACAATAATTGAAGGAGCAAGGACAGGAGAGATTGGTGATGGCAAGAT 526
Db 472 AGGAGTCATTGCCATGATAATTGAAGAGGCAAGAACTGGTGAATAGGTGATGGAAAGAT 531
QY 527 TTTTGTGTTGCTGTCAGATGTCATAAGAGTTAGGACAGCTGAGCGTGGGGAGAAAGC 586
Db 532 ATTCTTGACTCCCATCTCCGATGTTATAAGAGTTCCGCACTGGTGAACGGGGAGAAAGGC 591
QY 587 AGAGAGATGACTGGTG 603
Db 592 AGAGAGGATGATGGGAG 608

Search completed: May 28, 2004, 00:14:51
Job time : 1898.7 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 13:35:34 ; Search time 2651.35 Seconds
(without alignments)
14663.718 Million cell updates/sec

Title: US-09-756-541-14
Perfect score: 897
Sequence: 1 GCGGTGTCGGCGCTCTAGA.....AAAAAATACTCGAG 897

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl :

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rnd:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	897	100.0	897	6	AR125590	AR125590 Sequence
2	840	93.6	840	8	AF095454	AF095454 Ricinus c
3	594	66.2	594	6	AR125592	AR125592 Sequence
4	300.2	33.5	796	8	AY027892	AY027892 Medicago
5	286.4	31.9	817	6	AR125589	AR125589 Sequence
6	286.4	31.9	844	8	AF095455	AF095455 Arabidops
7	285	31.8	588	6	AR125591	AR125591 Sequence
8	285	31.8	591	6	AX507438	AX507438 Sequence
9	285	31.8	591	8	BT005209	BT005209 Arabidops
10	277.4	30.9	947	8	AY442185	AY442185 Lycopersi
11	217.2	24.2	902	8	AK068407	AK068407 Oryza sat
12	217.2	24.2	902	8	AK099152	AK099152 Oryza sat
13	215.4	24.0	1085	8	PPI489604	AJ489604 Pinus pin
14	107	11.9	686	1	AF017419	AF017419 Nostoc pu
15	103.2	11.5	2493	1	ASP251822	AJ251822 Anabaena
16	103.2	11.5	341880	1	AP003589	AP003589 Nostoc sp
17	99	11.0	384	1	FDGLNBPRT	X97327 F.diplosiph
18	98.8	11.0	191028	8	PPU38804	U38804 Porphyra pu
19	96.4	10.7	11978	1	U67574	U67574 Methanococc
20	96.4	10.7	110000	6	AR271569_01	Continuation [2 of
21	96.2	10.7	10338	1	U67464	U67464 Methanococc
22	96.2	10.7	110000	6	AR271569_14	Continuation (15 o
23	95.4	10.6	383	1	SYOGLNB	M62447 Synecococc
24	95.4	10.6	4348	1	AF079137	AF079137 Synecococ
25	94.4	10.5	452	1	SSGLMBP	X97496 Synecocyst
26	94.4	10.5	130001	1	D90915	D90915 Synecocyst
27	92	10.3	300143	1	AE017165	AE017165 Prochloro
28	90.4	10.1	164921	8	AF022186	AF022186 Cyanidium
29	90.4	10.1	299350	1	AP005370	AP005370 Thermosyn
30	88.4	9.9	9973	1	AE004295	AE004295 Vibrio ch
31	88.2	9.8	1622	1	AF120107	AF120107 Synecococ
32	86.8	9.7	300045	1	AE016803	AE016803 Vibrio vu
33	86.8	9.7	302320	1	AP005081	AP005081 Vibric pa
34	86.4	9.6	262202	1	EX572094	EX572094 Prochloro
35	85.2	9.5	247950	1	AP005340	AP005340 Vibrio vu
36	84.8	9.5	96109	6	AR408756	AR408756 Sequence
37	84.8	9.5	96109	6	AX067460	AX067460 Sequence
38	84	9.4	11371	1	AE006236	AE006236 Pasteurel
39	83.6	9.3	375	6	AR318761	AR318761 Sequence
40	83.6	9.3	298950	1	AP004597	AP004597 Oceanobac
41	83.4	9.3	9977	1	AE002548	AE002548 Neisseria
42	83.2	9.3	300880	1	AE016917	AE016917 Chromobac
43	82.2	9.2	300272	1	AE017213	AE017213 Geobacter
44	82	9.1	346792	1	EX571658	EX571658 Wolinella
45	81.8	9.1	349061	1	NMA222491	AL162753 Neisseria

ALIGNMENTS

RESULT 1
AR125590
LOCUS AR125590 897 bp DNA
DEFINITION Sequence 14 from patent US 6177275.
ACCESSION AR125590
VERSION AR125590.1 GI:14111652
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 897)
AUTHORS Coruzzi,G.M., Lam,H.-M. and Hsieh,M.-H.
TITLE Plant nitrogen regulatory P-PII genes
JOURNAL Patent: US 6177275-A 14 23-JAN-2001;
FEATURES Location/Qualifiers

linear PAT 16-MAY-2001

source		1. .897		/organism="unknown"		/mol_type="unassigned DNA"	
ORIGIN		Query Match		100.0%; Score 897; DB 6; Length 897;			
		Best Local Similarity		100.0%; Pred. No. 3.1e-211;			
		Matches 897; Conservative		0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	GGGGTGTGGCGCGCTCTAGAACTAGTGGATCCCGGGCTGCAGGAATTCGGCAGGCG	60				
Db	1	GGGGTGTGGCGCGCTCTAGAACTAGTGGATCCCGGGCTGCAGGAATTCGGCAGGCG	60				
QY	61	TACTGGAAACTGGGCTTGCTCACTCTCTTCATTCTAATAACATCAAGAAAGATTCCC	120				
Db	61	TACTGGAAACTGGGCTTGCTCACTCTCTTCATTCTAATAACATCAAGAAAGATTCCC	120				
QY	121	TGTTTGTGATTTTCAGTTTGTGTTTGTCCAGAGCTTAGACATTCCTGGTTTCTCATTAA	180				
Db	121	TGTTTGTGATTTTCAGTTTGTGTTTGTCCAGAGCTTAGACATTCCTGGTTTCTCATTAA	180				
QY	181	CACCGCGTCAAGCGCGTAAGATATGCCCGCTCGTTCCTGTGATTAATGCCCAAGCTC	240				
Db	181	CACCGCGTCAAGCGCGTAAGATATGCCCGCTCGTTCCTGTGATTAATGCCCAAGCTC	240				
QY	241	GCCTGACTACATTCCTGATGCTAAATCTACAAAGTGGAGCAATTCCTCAGGCCCTGGCG	300				
Db	241	GCCTGACTACATTCCTGATGCTAAATCTACAAAGTGGAGCAATTCCTCAGGCCCTGGCG	300				
QY	301	AGTCTCGCAAGTTTCTCGGCTTGTCTAATAAATGGTATTCGAGGTGTTACTGTTCTGA	360				
Db	301	AGTCTCGCAAGTTTCTCGGCTTGTCTAATAAATGGTATTCGAGGTGTTACTGTTCTGA	360				
QY	361	TGTTGAGGTTTGGTGCTCAAGGTGGTTCAACTGAGAGCGAGCGGCGCTCAGAAATTTTC	420				
Db	361	TGTTGAGGTTTGGTGCTCAAGGTGGTTCAACTGAGAGCGAGCGGCGCTCAGAAATTTTC	420				
QY	421	TGAAGCAAGTTTGTGCTAAGTTAAGTGAAGTGGAGATCGTGGTTAGCAAGACCAGTTGA	480				
Db	421	TGAAGCAAGTTTGTGCTAAGTTAAGTGGAGATCGTGGTTAGCAAGACCAGTTGA	480				
QY	481	GGATGTTATAGAAAAAATCATTGAGGAGCAAGAACTGGAGAGATTGGAGACGGCAAGAT	540				
Db	481	GGATGTTATAGAAAAAATCATTGAGGAGCAAGAACTGGAGAGATTGGAGACGGCAAGAT	540				
QY	541	TTTCTGTGCTGCTGTTTCAGATGTAATAAGAGTCGGCACTGGTGAGCGGGGTGATAAGGC	600				
Db	541	TTTCTGTGCTGCTGTTTCAGATGTAATAAGAGTCGGCACTGGTGAGCGGGGTGATAAGGC	600				
QY	601	TGAGAGGATGACAGGAGGGCGATCTGACATGAGTACTTCTGCTTGACTGTGACCCAGC	660				
Db	601	TGAGAGGATGACAGGAGGGCGATCTGACATGAGTACTTCTGCTTGACTGTGACCCAGC	660				
QY	661	AATATAGCATTCAGGACTAACTGTCCTTTGAGAAAGCCCGCCCTTATTAGCCATTATCC	720				
Db	661	AATATAGCATTCAGGACTAACTGTCCTTTGAGAAAGCCCGCCCTTATTAGCCATTATCC	720				
QY	721	AGTATAGCTTGATAATTGAAATTTTGTGTTTCTTAACTAAAGAAACAAAGATCTTTTCA	780				
Db	721	AGTATAGCTTGATAATTGAAATTTTGTGTTTCTTAACTAAAGAAACAAAGATCTTTTCA	780				
QY	781	TTATCCTGTTGATGATAATTGAAAACGGAGGATCGCGAATTTGTTCAAGTCTTGCAAG	840				
Db	781	TTATCCTGTTGATGATAATTGAAAACGGAGGATCGCGAATTTGTTCAAGTCTTGCAAG	840				
QY	841	ATAAATAACAAGAGAGGAGTAAATGTTAACAAAAAAATAAAAAAATAAAAACTCGAG	897				
Db	841	ATAAATAACAAGAGAGGAGTAAATGTTAACAAAAAAATAAAAAAATAAAAACTCGAG	897				
RESULT 2		AF095454		840 bp		mrna	
LOCUS				linear		PLN 19-NOV-1998	

DEFINITION		Ricinus communis PII protein mRNA, partial cds.	
ACCESSION		AF095454	
VERSION		AF095454.1 GI:3885940	
KEYWORDS		Ricinus communis (castor bean)	
SOURCE		Ricinus communis	
ORGANISM		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Malpighiales; Euphorbiaceae; Acalyphoideae; Acalyphaeae; Ricinus.	
REFERENCE		1 (bases 1 to 840)	
AUTHORS		Hsieh, M.H., Lam, H.M., van de Loo, F.J. and Coruzzi, G.	
TITLE		A PII-like protein in Arabidopsis: putative role in nitrogen sensing	
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 95 (23), 13965-13970 (1998)	
MEDLINE		99030678	
PUBMED		9811909	
REFERENCE		2 (bases 1 to 840)	
AUTHORS		Hsieh, M.H., Lam, H.M., van de Loo, F.J. and Coruzzi, G.M.	
TITLE		Direct Submission	
JOURNAL		Submitted (29-SEP-1998) Biology, New York University, 100 Washington Square East, New York, NY 10003, USA	
FEATURES		Location/Qualifiers	
source		1..840	
		/organism="Ricinus communis"	
		/mol_type="mrna"	
		/db_xref="taxon:3988"	
CDS		<1..595	
		/function="putative role in nitrogen sensing"	
		/note="GLB1"	
		/codon_start=2	
		/product="PII protein"	
		/protein_id="AAC78332.1"	
		/db_xref="GI:3885941"	
		/translation="HEATKLGLLTPLHSNNIKKFFVDFSLFCEPLRHSRFSHENTAVKRVYAPVVPVINAQSSPDYIPDAKFKVEAILRPWRVSQVSSALLKIGIRGVTVSDVRGFGAQQGSTERQGSFSEDFVAKVMEIVVSKDQVEDVIEKIIIEEARTGEIGD	
		GKIFLLPVSDVIRVTRGGRDKAERMTGGRSDMSTSA"	
ORIGIN		Query Match	
		Best Local Similarity	
		Matches 840; Conservative	
		0; Mismatches 0; Indels 0; Gaps 0;	
QY	52	GCACGAGGCTACTCGAAACTGGGCTTGCTCACTCCTCTTCAATCTAATAACATCAAGAA	111
Db	1	GCACGAGGCTACTCGAAACTGGGCTTGCTCACTCCTCTTCAATCTAATAACATCAAGAA	60
QY	112	AGAAATCCCTGTTTGTGATTTCAGTTTGTGTTTGTCCAGAGCTTAGACATTCCTGGTTTC	171
Db	61	AGAAATCCCTGTTTGTGATTTCAGTTTGTGTTTGTCCAGAGCTTAGACATTCCTGGTTTC	120
QY	172	TCACTTTAAACACCGCGGTCAAGCGCGTAAGATATGCCCGTGGTTCCTGTGATTAATGC	231
Db	121	TCACTTTAAACACCGCGGTCAAGCGCGTAAGATATGCCCGTGGTTCCTGTGATTAATGC	180
QY	232	CCAAAGCTCGCTGACTACATTCCTGATGCTAAATCTACAAAGTGGAGCAATTCCTCAG	291
Db	181	CCAAAGCTCGCTGACTACATTCCTGATGCTAAATCTACAAAGTGGAGCAATTCCTCAG	240
QY	292	GCCCTGGCGAGTCTCGCAAGTTTCTCGGCTTGTCTAATAATGGTATTCGAGGTGTAC	351
Db	241	GCCCTGGCGAGTCTCGCAAGTTTCTCGGCTTGTCTAATAATGGTATTCGAGGTGTAC	300
QY	352	TGTTTCTGATGTTTCGAGGTTTGTGCTCAAGGTGGTTCAACTGAGAGGCGGCGGCTC	411
Db	301	TGTTTCTGATGTTTCGAGGTTTGTGCTCAAGGTGGTTCAACTGAGAGGCGGCGGCTC	360
QY	412	AGAAATTTCTGAAGACAAAGTTTGTGTTGCTAAAAGTTAAGATGGAGATCGTGGTAGCAAGA	471
Db	361	AGAAATTTCTGAAGACAAAGTTTGTGTTGCTAAAAGTTAAGATGGAGATCGTGGTAGCAAGA	420
QY	472	CCAGGTTGAGGATGTTATAGAAAAAATCATTTGAGGAGGCAAGAACTGGAGAGATTGGAGA	531

Db 421 CCAGGTTGAGGATGTTATAGAAAAAATCAITGAGGAGCGCAAGAACTGGAGAGATTGGAGA 480
QY 532 CGGCAAGATTTTCTTGTCCTGTTTTCAGATGTAATAAGAGTCGCACTGGTGAGCGGGG 591
Db 481 CGGCAAGATTTTCTTGTCCTGTTTTCAGATGTAATAAGAGTCGCACTGGTGAGCGGGG 540
QY 592 TGATAAGGCTGAGAGGATGACAGGAGGGCGATCTGACATGAGTACTTCTGCTTGACTGCT 651
Db 541 TGATAAGGCTGAGAGGATGACAGGAGGGCGATCTGACATGAGTACTTCTGCTTGACTGCT 600
QY 652 GTGACCAGCAATATAGCAATTCAGGACTAACTGTCTTTTGAGAAAGCCCGCCCTTATTAG 711
Db 601 GTGACCAGCAATATAGCAATTCAGGACTAACTGTCTTTTGAGAAAGCCCGCCCTTATTAG 660
QY 712 CCATTATCCAGTATAGCTTGATAAATTTGAATTTTGTGTTTCTTAACAAAGAAACAAAG 771
Db 661 CCATTATCCAGTATAGCTTGATAAATTTGAATTTTGTGTTTCTTAACAAAGAAACAAAG 720
QY 772 ATCTTTTCAATTATCCCTGTTGATGATAAATTGAAACCGGAAGGATCGCGAATTTGTCAAAGT 831
Db 721 ATCTTTTCAATTATCCCTGTTGATGATAAATTGAAACCGGAAGGATCGCGAATTTGTCAAAGT 780
QY 832 GCTTGCAAGATAAATAACAAGAGGAGGAGTAATGTTAAACAAAAAATAAAAAA 891
Db 781 GCTTGCAAGATAAATAACAAGAGGAGGAGTAATGTTAAACAAAAAATAAAAAA 840

RESULT 3
AR125592 LOCUS AR125592 594 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 16 from patent US 6177275.
ACCESSION AR125592
VERSION AR125592.1 GI:14111654
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 594)
AUTHORS Coruzzi,G.M., Lam,H.-M. and Hsieh,M.-H.
TITLE Plant nitrogen regulatory P-PII genes
JOURNAL Patent: US 6177275-A 16 23-JAN-2001;
FEATURES Location/Qualifiers
source 1..594
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 66.2%; Score 594; DB 6; Length 594;
Best Local Similarity 100.0%; Pred. No. 3.2e-136;
Matches 594; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 CGGCACGAGGCTACTCGAAACTGGGCTTGCTCACTCCTCTTCATTCTAATAACATCAAG 109
Db 1 CGGCACGAGGCTACTCGAAACTGGGCTTGCTCACTCCTCTTCATTCTAATAACATCAAG 60
QY 110 AAAGAATCCCTGTTTGTGATTTCAGTTTGTGTTGTCAGAGCTTAGACATTCTCGGTTT 169
Db 61 AAAGAATCCCTGTTTGTGATTTCAGTTTGTGTTGTCAGAGCTTAGACATTCTCGGTTT 120
QY 170 TCTCATTTTAACACCGCGGTCAAGCGCGTAAGATATGCCCCCGTTCCTGTGATTAAT 229
Db 121 TCTCATTTTAACACCGCGGTCAAGCGCGTAAGATATGCCCCCGTTCCTGTGATTAAT 180
QY 230 GCCCAAAGCTCGCCTGACTACATTCTCTGATGCTAAATTCTACAAAGTGGAGCAATTCTC 289
Db 181 GCCCAAAGCTCGCCTGACTACATTCTCTGATGCTAAATTCTACAAAGTGGAGCAATTCTC 240
QY 290 AGGCCCTGGGAGTCTCGCAAGTTTCCTCGGCTTTGCTAAAAAATGGTATTCGAGGTGTT 349
Db 241 AGGCCCTGGGAGTCTCGCAAGTTTCCTCGGCTTTGCTAAAAAATGGTATTCGAGGTGTT 300
QY 350 ACTGTTTCTGATGTTGAGGTTTGTGGTGCTCAAGGTGGTTCAACTGAGAGGCGAGCGGC 409

Db 301 ACTGTTTCTGATGTTTCGAGGTTTGGTGCTCAAGTGGTTCAACTGAGAGCGAGCGGC 360
QY 410 TCAGAATTTTCTGAAGACAAGTTTGTGTTGCTAAAGTTAAGATGGAGATCGTGGTTAGCAA 469
Db 361 TCAGAATTTTCTGAAGACAAGTTTGTGTTGCTAAAGTTAAGATGGAGATCGTGGTTAGCAA 420
QY 470 GACCAGGTTGAGGATGTTATAGAAAAAATCAATTGAGGAGGCAAGAACTGGAGAGATTGGA 529
Db 421 GACCAGGTTGAGGATGTTATAGAAAAAATCAATTGAGGAGGCAAGAACTGGAGAGATTGGA 480
QY 530 GACGGCAAGATTTTCTTGCTGCTGCTGTTTCAGATGTAATAAGAGTCCGCACTGGTGAGCGG 589
Db 481 GACGGCAAGATTTTCTTGCTGCTGCTGTTTCAGATGTAATAAGAGTCCGCACTGGTGAGCGG 540
QY 590 GGTGATAAGGCTGAGAGGATGACAGGAGGGCGGATCTGACATGAGTACTTCTGCT 643
Db 541 GGTGATAAGGCTGAGAGGATGACAGGAGGGCGGATCTGACATGAGTACTTCTGCT 594

RESULT 4
AY027892 LOCUS AY027892 796 bp mRNA linear PLN 12-MAR-2001
DEFINITION Medicago sativa PII protein (GLNB) mRNA, complete cds.
ACCESSION AY027892
VERSION AY027892.1 GI:13277514
KEYWORDS
SOURCE Medicago sativa
ORGANISM Medicago sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 796)
AUTHORS Garcia-Ibáñez,D. and Sengupta-Gopalan,C.
TITLE Characterization of PII (GLNB) in alfalfa
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 796)
AUTHORS Garcia-Ibáñez,D. and Sengupta-Gopalan,C.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-2001) Molecular Biology Program, New Mexico State University, Corner of Knox and College, Las Cruces, NM 88003, USA
FEATURES Location/Qualifiers
source 1..796
/organism="Medicago sativa"
/mol_type="mRNA"
/db_xref="taxon:3879"
gene 1..796
/gene="GLNB"
CDS 1..585
/gene="GLNB"
/codon_start=1
/product="PII protein"
/protein_id="AAK16221.1"
/db_xref="GI:13277515"
/translation="MTLIAKPNVFNLFNINETQIPFSSFSVIRKFRFGDSSHRNVVL
RSMGNASVLPKIRAQNLPDYVPESKFYKVEAILRPWRIPOVSSGLLKMIGIRGTVSDV
KGFQAQGSKEKRGSEFSEDFNFAKVKMEIVVRKDQVEAVINKIMETARTGEIGDGK
IFLIPVSDVIRITGERGEQAERMAGGLTDALV"

ORIGIN

Query Match 33.5%; Score 300.2; DB 8; Length 796;
Best Local Similarity 72.0%; Pred. No. 1.8e-63;
Matches 408; Conservative 0; Mismatches 153; Indels 6; Gaps 1;
QY 62 ACTGCGAAACTGGGCTTGCTCACTCCTCTTCATTCTAATAACATCAAGAAAGAAATCCCT 121
Db 10 ATTGCGAAACCGAACGTTTCAACGGTTTGAATTTTAAATATCAATGAACCTCAAATTCCT 69
QY 122 GTTTTGTGATTTTCAGTTTCTTTTGTCCAGAGCTTAGACATTCTCGGTTTCTCACCTTAC 181
Db 70 TTTTCAGCTTCAGCGGTCAATTCGCAAGCGTTTCGGAGATTCTCTCATCGCAAT 123


```
QY 45 GAATTCGGACGAGGCTACTCGGAACTGGGCTTGCTCACTCCTCTTCATTCTAATAACA 104
    |||||
Db 26 GAATCATGGCGCGTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTCTATTCTG 85

QY 105 TCAAGAAAGAAATCCCTGTTTTGATTTCAGTTTGTGTTTGTCCAGAGCTTAGACATTCTC 164
    |||||
Db 86 ATCGAAAGAAACAATGCTTTCTCTGATTGCAATTGCAATTTGTTCTGGAATCAGACATTCCC 145

QY 165 GGTTTTTCACATTTAAACACCGCGGTCAAGCGGTAAGATATGCC---CCGTCGTTCTCTG 221
    |||||
Db 146 GACCAATCTTGCCCTCGAATTGGTTCACAAAGTCACCGAGTAATAACAGTCGTTGTTTACCTG 205

QY 222 TGATTAAATGCCAAAGCTCGCCTGACTACATTCCTGTGATGCTAAATTCACAAAGTGAAG 281
    |||||
Db 206 TCGTTAGTGCCCAATATCTTCTGATTATATCCAGACTCGAAATTTTACAAAGGTGAAG 265

QY 282 CAATTCTCAGGCCCTGGCGAGTCTCGCAAGTTTCTCGGCTTTGCTAAATAATGTTATTC 341
    |||||
Db 266 CAATTGTCAGACCATGGAGATCCAGCAAGTTTCATCGGCTTTACTGAAATCGGGATTTC 325

QY 342 GAGGTGTTACTGTTTCTGATGTTTCGAGGTTTTCGCTCAAGGTGGTTCACACTGAGAGGC 401
    |||||
Db 326 GAGGTGTTACTGTTTCTGATGTGAGAGGTTTGGTGCAAGGAGGTTCTACCGAGAGAC 385

QY 402 AGGCGGCTCAGAAATTTTCTGAAGACAAGTTTGTGCTTAAAGTTAAGATGAGATCGTGG 461
    |||||
Db 386 ACGGTGGCTCTGAGTCTCGGAAGACAAATTTGTGCTAAAGTTAAGATGAGAAATCGTTG 445

QY 462 TTAGCAAAGACCAAGTTGAGGATGTTATAGAAAAAATCATTGAGGAGGCAAGAACTGGAG 521
    |||||
Db 446 TTAAGAAAGACCAAGTGAATCTGTAATCAACACAATAATTGAAGGAGCAAGACAGGAG 505

QY 522 AGATTGGAGACGCAAGATTTTCTTGCTGCTGTTTTCAGATGTAATAAGAGTCCGCACTG 581
    |||||
Db 506 AGATTGGTGATGCGCAAGATTTTGTGTTTGCCTGTGTCAGATGTCTAAGAGTTAGGACAG 565

QY 582 GTGAGCGGGTGATAGGCTGAGAGGATGACAGGAG 617
    |||||
Db 566 GTGAGCGGTGGGAGAAAGCAGAGAGATGACTGGTG 601

RESULT 7
AR125591
LOCUS AR125591 588 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 15 from patent US 6177275.
ACCESSION AR125591
VERSION AR125591.1 GI:14111653
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 588)
AUTHORS Coruzzi,G.M., Lam,H.-M. and Hsieh,M.-H.
TITLE Plant nitrogen regulatory P-PII genes
JOURNAL Patent: US 6177275-A 15 23-JAN-2001;
FEATURES
    source
        Location/Qualifiers
            1..588
                /organism="unknown"
                /mol_type="unassigned DNA"

ORIGIN
Query Match 31.8%; Score 285; DB 6; Length 588;
Best Local Similarity 70.3%; Pred. No. 1e-59;
Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

QY 56 GAGGCTACTCGGAACCTGGGCTTGCTCACTCCTCTTCATTCTAATAACATCAAGAAAGAA 115
    |||||
Db 7 GCGTCAATGACGAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGAAAGAAC 66

QY 116 TTCCTGTTTTGATTCAGTTTGTGTTTGTCCAGAGCTTAGACATTCGCGTTTCTCAC 175
    |||||
Db 67 ATTGCTTTCCTGATTCGATTTGATTTGTTCTGGAATTCAGACATTCGCCACCATCTTGC 126
```

```
QY 176 TTTAACACCGCGGTCAAGCGGTAAGATATGCC---CCCGTCGTTCTGTGATTAATGCC 232
    |||||
Db 127 CTCGATTTGGTCACAAAGTCACCGAGTAATAACAGTCGTTGTTTACCTGTCTAGTGCC 186

QY 233 CAAAGCTCGCCTGACTACATTCTCTGATGCTAAATTTCTACAAAGTGAAGCAATTTCTCAGG 292
    |||||
Db 187 CAAATATCTTCTGATTATATCCAGACTCGAAATTTTACAAAGTGAAGCAATTTGTTCAGA 246

QY 293 CCCTGGCGAGTCTCGCAAGTTTCTCGGCTTTGCTAAATAATGTTATTCGAGGTGTTACT 352
    |||||
Db 247 CCATGGAGAAATCCAGCAAGTTTTCATCGGCTTTACTGAAATCGGATTCGAGGTGTTACT 306

QY 353 GTTCTGTGATTCGAGGTTTGGTGTCTCAAGGTGGTTCAACTGAGAGGCGGGGGCTCA 412
    |||||
Db 307 GTTCTGTGATGAGAGGTTTGGTGTCAACAGGAGGTTCTACCGAGAGACACGTTGGCTCT 366

QY 413 GAAATTTTCTGAAGACAAAGTTTGTGCTAAAGTTAAGATGAGATCGTGGTTAGCAAAGAC 472
    |||||
Db 367 GAGTTCTCGGAAGACAAATTTGTGCTAAAGTTAAGATGGAATCGTTGTTAAGAAAGAC 426

QY 473 CAGGTTGAGGATGTTATAGAAAAAATCATTGAGGAGGCAAGAACTGGAGAGATTGGAGAC 532
    |||||
Db 427 CAAGTGAATCTGTAATCAACACAATAATTGAAGGAGCAAGACAGGAGAGATTGGTGAT 486

QY 533 GGCAAGATTTTCTTGCTGCTGTTTTCAGATGTAATAAGACTCCGCACTGGTGAGCGGGCT 592
    |||||
Db 487 GGCAAGATTTTGTGCTGCTGCTCAGATGTCATAAGAGTTAGGACAGGTGAGCGGTGG 546

QY 593 GATAAGGCTGAGAGGATGACAGGAG 617
    |||||
Db 547 GAGAAAGCAGAGAGATGACTGGTG 571

RESULT 8
AX507438 591 bp DNA linear PAT 27-SEP-2002
LOCUS AX507438
DEFINITION Sequence 2133 from Patent WO0216655.
ACCESSION AX507438
VERSION AX507438.1 GI:23388675
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Harper,J.F., Kreps,J., Wang,X. and Zhu,T.
TITLE Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
JOURNAL Patent: WO 0216655-A 2133 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)
FEATURES
    source
        Location/Qualifiers
            1..591
                /organism="Arabidopsis thaliana"
                /mol_type="unassigned DNA"
                /db_xref="taxon:3702"

ORIGIN
Query Match 31.8%; Score 285; DB 6; Length 591;
Best Local Similarity 70.3%; Pred. No. 1e-59;
Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

QY 56 GAGGCTACTCGGAACCTGGGCTTGCTCACTCCTCTTCATTCTAATAACATCAAGAAAGAA 115
    |||||
Db 7 GCGTCAATGACGAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGAAAGAAC 66

QY 116 TTCCTGTTTTGATTCAGTTTGTGTTTGTCCAGAGCTTAGACATTCGCGTTTCTCAC 175
    |||||
Db 67 ATTGCTTTCCTGATTCGATTTGATTTGTTCTGGAATTCAGACATTCGCCACCATCTTGC 126

QY 176 TTTAACACCGCGGTCAAGCGGTAAGATATGCC---CCCGTCGTTCTGTGATTAATGCC 232
    |||||
```


Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica rice.

COMMENT

URL : <http://cdna01.dna.affrc.go.jp/cdna/>
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.

FEATURES

source

1. 902
/organism="Oryza sativa (japonica cultivar-group)"
/mol type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J013149B08"

ORIGIN

Query Match 24.2%; Score 217.2; DB 8; Length 902;
Best Local Similarity 69.7%; Pred. No. 6.3e-43;
Matches 294; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Qy	238	CTCGCTGACTACATTCCTGATGCTTAATCTACAAGTGGAGCAATTCTCAGGCCCTG	297
Db	267	CGACGACGGTACCAGCCGGAGTCGGAGTCTCAAGGTGGAGGCAATCTGAGGCCATG	326
Qy	298	GCGAGTCTCGCAAGTTTCCTCGGCTTTGCTAAAGTTAAGATGGAGATCGTGTACTGTTCT	357
Db	327	GAGGGTGCCTTATGTGTATCATCGGGTTTGCTGCAATGGGGATCAGAGCGGTGACGGTGC	386
Qy	358	TGATGTTGAGGTTTGGTGTCTCAAGTGGTTCAACTGAGAGCGCAGGCGGCTCAGAATT	417
Db	387	CGACGTGCGGGTTTCGGCGCACAGGCGGGTCAACTGAGAGGCGATGAAGGGTCAGAATT	446
Qy	418	TTCTGAAGACAAGTTTGTGTAAAGTTAAGATGGAGATCGTGGTTAGCAAGACCAGGT	477
Db	447	TGCAGAGATACATTTATTGTATAAGTTAAGATGGAATAGTGGTGTCCAAGGATCAGGT	506
Qy	478	TGAGGATGTTATAGAAAAATCATTGAGGAGCGCAAGAACTGGAGAGATTGGAGACGGCAA	537
Db	507	TGAAGCTGTTGTTGACAAGATAATTGAAAAGCAAGAACAGGAGAAATTGGTATGGAAA	566
Qy	538	GATTTCTTGCTGCCTGTTTCAGATGTATAAGAGTCGGCACTGGTCAGCGGGGTGATAA	597
Db	567	AATATTTTGATACCCGTGTCGGACCGTATCAGAATACGCACCGCGCAACGAGGGGAGCG	626
Qy	598	GGCTGAGAGATGACAGGAGGGCGATCTGACATGAGTACTTCTGCTTACTGCTGTGACC	657
Db	627	AGCGGAGAGGATGGCCGGAGGGCTGGCGGACAAGCTGTCTCTCAGCAATGCCGATCTCATG	686
Qy	658	AG 659	

Db 687 AG 688

RESULT 12

AK099152
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) cDNA clone:J023070H02, full insert sequence.
AK099152
ACCESSION
VERSION
AK099152.1 GI:32984361
KEYWORDS
SOURCE
FLI_CDNA; CAP trapper.
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE

AUTHORS

The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., Arakawa,T., Fukuda,S., Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Matsubara,K., RIKEN: Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.

TITLE

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
Science 301 (5631), 376-379 (2003)
22752273
12869764

REFERENCE

AUTHORS

2 (bases 1 to 902)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Kagawa,I., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kawamata,M., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.

Direct Submission

JOURNAL

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT

This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : <http://cdna01.dna.affrc.go.jp/cdna/>
NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,

Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.

FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,M., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.

FEATURES
source
1..902
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023070H02"

ORIGIN

Query Match 24.2%; Score 217.2; DB 8; Length 902;
Best Local Similarity 69.7%; Pred. No. 6.3e-43;
Matches 294; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 238 CTCGCCCTGACTACATTCCTGATGCTAAATCTACAAAGTGAAGCAATCTCAGGCCCTG 297
Db 267 CCAGCAGGGTACCAGCCGGAGTCGGAGTTCTACAGGTGAGGCAATCCTGAGGCCATG 326
QY 298 GCGAGTCTCGCAAGTTTCCTCGGCTTTGCTTAAAAATTGGTATTCGAGGTGTTACTGTTTC 357
Db 327 GAGGGTGCCTTATGTGTCATCGGGTTTGCTGCAATGGGGATCAGAGCGGTGACGGTGTC 386
QY 358 TGATGTTTCGAGGTTTTGCTGCTCAAGGTGGTTCAACTGAGAGCGCAGCGGCTCAGAATT 417
Db 387 CGACGTGCGGGGTTTTCGCGCACACAGGCGGSGTCAACTGAGAGGCATGAAGGGTCAGAATT 446
QY 418 TTCTGAAGACAAGTTTGTGCTTAAAGTTAAGATGGAGATCGTGGTTAGCAAAGACCAGGT 477
Db 447 TGCAGAAGATACATTTATTGATAAAGTTAAGATTGAATAGTAATAGTGGTGTCTCCAAGGATCAGT 506
QY 478 TGAGGATGTTATAGAAAAATCATTGAGGAGCGCAAGAACTGGAGAGATTGGAGACGGCAA 537
Db 507 TGAAGCTGTTGTTGACACAGATAATTGMAAAGGCAAGAACAGAGAGAAATTGGTGATGAAA 566
QY 538 GATTTTCTGCTGCCTGTTTCAGATGTAATAAGAGTCCGCACTGGTGAGCGGGTGATAA 597
Db 567 AATATTTTGTATACCCGTGTCGGACGTGATCAGATAACGACCGGGAACGAGGGAGCG 626
QY 598 GGCTGAGAGGATGACAGAGGGCGATCTGCATGAGTACTTCTGCTTGACTGCTGTGACC 657
Db 627 ACCGGAGAGGATGGCCGAGGGCTGGCGGACAAGCTGTCTCAGCAATGCCGATCTCATG 686
QY 658 AG 659
Db 687 AG 688

RESULT 13
PPI489604 1085 bp mRNA linear PLN 11-OCT-2003
LOCUS PPI489604
DEFINITION Pinus pinaster mRNA for PII-like protein (glb gene).

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

FEATURES
source

gene
CDS

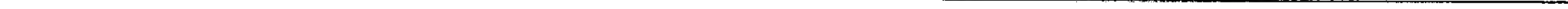
ORIGIN

Query Match 24.0%; Score 215.4; DB 8; Length 1085;
Best Local Similarity 73.2%; Pred. No. 1.8e-42;
Matches 276; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 242 CCTGACTACATTCCTGATGCTAAATCTCAAAAGTGAAGCAATCTCAGGCCCTGGCGA 301
Db 327 CCAGATTATGTCCAGAAAGCCAACTTTACAAAGTAGAAGCAATATGAGGCCATGGCGC 386
QY 302 GTCTCGCAAGTTTCCTCGGCTTTGCTTAAAAAATTGGTATTCGAGGTGTTACTGTTCTGAT 361
Db 387 ATCTCCCATGTGACTACGGGTCTATTGAAAAATGGGGATTCTGTGGCTAACTGTCTCTGAT 446
QY 362 GTTCGAGGTTTTGGTGCTCAAGGTGGTTCAACTGAGAGCGAGCGGCTCAGAAATTTTCT 421
Db 447 GTTAGAGGTTTTGGAGTTTCAGGCTGGATCTGCAGAACGGCAAGCAGGCTCCGAGTTTTCT 506
QY 422 GAAGACAAGTTTGTTCCTAAAGTTAAGATGGAGATCGTGTGTTAGCAAAGACCAGGTTGAG 481
Db 507 AAAGACAATTTGTGTCAAAGTAAAGATGGAGATTGTGGTATCTAAAGATCAGGTAGAA 566
QY 482 GATGTTATAGAAAAAATCAATTCAGGAGGGCAAGAACTGGAGAGATTGGAGACGGCAAGATT 541
Db 567 GCAGTAATTGATGCAATCATTTGATGAGGCAAGAACTGGAGAAATTGGAGATGGCAAAATA 626
QY 542 TTCTTGCTGCTGTTTCAGATGTAATAAGATCCGCACTGGTGAGCGGGGTGATAAGGCT 601
Db 627 TTGTGGTTCAGTTGCAGATGTCATTCTGTGTGAGAAACAGGTGAGCGGGGACTTGAAGCCA 686
QY 602 GAGAGGATGACAGGAGG 618
Db 687 GAGAGAAATGGCTGGTG 703

AJ489604
AJ489604.1 GI:37653226
glb gene; PII-like protein.
Pinus pinaster
Pinus pinaster
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
1
Canton,F.J.
A PII-like protein from xylem of adult Maritime pine trees
Unpublished
2 (bases 1 to 1085)
Canton,F.J.
Direct Submission
Submitted (11-JUN-2002) Canton F.J., Molecular Biology and Biochemistry, Faculty of Sciences, Campus de Teatinos s/n, E-29071, SPAIN
Location/Qualifiers
1..1085
/organism="Pinus pinaster"
/mol_type="mRNA"
/db_xref="taxon:71647"
/tissue_type="xylem"
/dev_stage="adult tree"
1..1085
/gene="glb"
42..773
/gene="glb"
/function="putative nitrogen sensor protein"
/codon_start=1
/product="PII-like protein"
/protein_id="CAD33967.1"
/db_xref="GI:37653227"
/translation="MAAHLPLFVKGSIFSLPSSSTSSCTYSISHTAASPGFHSPP
AAANKSVKLGTMQMDGVRSVKHRMTSLHAKMEKROQIKASQAQPNGTNTHPDYVPEA
NFIKVEAILRPWIRISHVTTGLLKMGIKRGVTSDVRGFGVQAGSABRQAGSEFSKDNFV
SKVKMEIIVSKDQVEAVIDAIIIDBARTGEIGDKIFVVPVADVIRVRTGERDLKPERM
AGDQRYLQVYIKRLQATKSLAL"

Search completed: May 27, 2004, 22:22:42
Job time : 2654.6 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 13:28:59 ; Search time 312.835 Seconds
(without alignments)
12180.972 Million cell updates/sec

Title: US-09-756-541-14
Perfect score: 897
Sequence: 1 GCGGTGTCGGCGCTCTAGA.....AAAAAATCGAG 897

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.*
1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002s.*
7: geneseqn2003as.*
8: geneseqn2003bs.*
9: geneseqn2003cs.*
10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	897	100.0	897	4	Aaf58582 Ricinus c
2	594	66.2	594	4	Aaf58584 Ricinus c
3	286.4	31.9	817	4	Aaf58581 Arabidops
4	285	31.8	588	4	Aaf58583 Arabidops
5	285	31.8	591	6	Abz14328 Arabidops
6	96.4	10.7	110000	2	Continuation (2 of
7	96.2	10.7	110000	2	Continuation (15 o
8	84.8	9.5	96109	4	Aaf28548 Genomic f
9	83.6	9.3	375	8	Ada30024 DNA encod
10	81.4	9.1	339	7	Acf72027 Photorhab
11	81.4	9.1	110000	7	Continuation (53 o
12	81.4	9.1	110000	7	Continuation (4 of
13	77.6	8.7	110000	2	Continuation (4 of
14	77	8.6	363	7	Abz40022
15	73	8.1	16526	3	Aaa81472 N. mening
16	73	8.1	110000	3	Continuation (7 of
17	73	8.1	172325	3	Aaf21613 Neisseria
18	73	8.1	349980	3	Aaf21612 Neisseria
19	70.8	7.9	336	5	Aah68445 C glutami
20	70.8	7.9	417	5	Aah67222 C glutami
21	70.8	7.9	349980	5	Aah68531 C glutami
22	69	7.7	1015	6	Abn74670 Bovine em
23	68.8	7.7	999	6	Abn74679 Bovine em

24	65.2	7.3	770	5	ABA13777	Abal3777 Human ner
25	64.8	7.2	290	4	AAF71296	Aaf71296 Corynebac
26	64.6	7.2	654	4	AAH34361	Aah34361 Human col
27	64.2	7.2	1164	5	AAS88441	Aas88441 DNA encod
28	64.2	7.2	11095	4	AAS46256	Aas46256 DNA encod
29	64	7.1	939	6	ABN74051	Abn74051 Bovine em
30	63.8	7.1	845	5	AAD05422	Aad05422 Human gen
31	63.8	7.1	845	7	ACC50696	Acc50696 Human sec
32	63.8	7.1	845	7	ABZ71380	Abz71380 Secreted
33	63.8	7.1	845	8	ADB91328	Adb91328 Human sec
34	63.8	7.1	845	9	ADC73801	Adc73801 Human sec
35	63.6	7.1	958	6	ABN74057	Abn74057 Bovine em
36	63.6	7.1	998	6	ABN74047	Abn74047 Bovine em
37	63.4	7.1	973	6	ABN74320	Abn74320 Bovine em
38	63.2	7.0	404	6	ABL69195	Ab169195 Prostate
39	63.2	7.0	404	6	ABK64407	Abk64407 Human ben
40	62.8	7.0	882	3	AAC63454	Aac63454 Human sec
41	62.8	7.0	882	7	ABZ73574	Abz73574 Secreted
42	62.8	7.0	882	7	ADA98087	Ada98087 Human sec
43	62.8	7.0	1012	4	ABA77106	AbA77106 Prolifera
44	62.8	7.0	2163	4	AAF91900	Aaf91900 Human sec
45	62.6	7.0	570	4	AAH33656	Aah33656 Human col

ALIGNMENTS

RESULT 1
AAF58582
ID AAF58582 standard; cDNA; 897 BP.
XX
AC AAF58582;
XX
DT 23-APR-2001 (first entry)
XX
DE Ricinus communis P-PII cDNA.
XX
KW Castor bean; PII; plant nitrogen regulatory gene; P-PII;
KW nitrogen assimilation; transgenic plant; herbicide screening; ss.
XX
OS Ricinus communis.
XX
PN US6177275-B1.
XX
PD 23-JAN-2001.
XX
PF 23-JUL-1997; 97US-00899330.
XX
PR 24-JUL-1996; 96US-0022328P.
XX
PA (UTNY) UNIV NEW YORK STATE.
XX
PI Coruzzi GM, Lam H, Hsieh M;
XX
DR WPI; 2001-158572/16.
DR P-PSDB; AAB69496.
XX
PT Novel P-PII genes capable of regulating plant nitrogen assimilation,
PT useful for transgenic plant production, and as probes for isolating
PT additional genomic clones having P-PII gene promoters.
XX
PS Claim 2; Fig 13; 35pp; English.

The present sequence encodes a nitrogen regulatory PII protein. Novel plant PII (also called P-PII) nucleotide sequences have been isolated. They are useful for regulating nitrogen assimilation in plants, and in transgenic plant production. They are also used to engineer organisms that overexpress wild-type or mutant P-PII regulatory proteins. P-PII proteins are useful for in vitro screening of herbicides. P-PII nucleotides may be used as probes for isolating additional genomic clones with the promoters of P-PII genes. P-PII promoters are light- and/or sucrose-inducible, and are suitable for genetic engineering of plants

```
SQ      Sequence 897 BP; 254 A; 171 C; 218 G; 254 T; 0 U; 0 Other;
      Query Match      100.0%; Score 897; DB 4; Length 897;
      Best Local Similarity 100.0%; Pred. No. 5.8e-212;
      Matches 897; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 GCGGTGTCGGCGCTCTAGAACTAGTGGATCCCGCGGCTGCAGGAATTCGGCACGAGGC 60
Db      1 GCGGTGTCGGCGCTCTAGAACTAGTGGATCCCGCGGCTGCAGGAATTCGGCACGAGGC 60
QY      61 TACTGCGAACTGGGCTTGCTCAGTCTCTCTCATTTCTAATAACATCAAGAAAGAAATTCCTCC 120
Db      61 TACTGCGAACTGGGCTTGCTCAGTCTCTCTCATTTCTAATAACATCAAGAAAGAAATTCCTCC 120
QY      121 TGTCTTTGATTTTCAAGTTTGTCTTGTCCAGAGCTTAGACATTTCTCGGTTTCTCATTAA 180
Db      121 TGTCTTTGATTTTCAAGTTTGTCTTGTCCAGAGCTTAGACATTTCTCGGTTTCTCATTAA 180
QY      181 CACCGCGGTCAAGCGGTAAGATATGCCCGCGTCTGTTCTGTGATTATGCCCCAAAGCTC 240
Db      181 CACCGCGGTCAAGCGGTAAGATATGCCCGCGTCTGTTCTGTGATTATGCCCCAAAGCTC 240
QY      241 GCCTGACTACATTCCTGATGCTAAATTTCTACAAAGTGAAGCAATTTCTCAGGCCCTGGCG 300
Db      241 GCCTGACTACATTCCTGATGCTAAATTTCTACAAAGTGAAGCAATTTCTCAGGCCCTGGCG 300
QY      301 AGCTCGCAAGTTTCTCGGCTTTGCTAAATAATGTTATTCGAGGTGTTACTGTTCTGA 360
Db      301 AGCTCGCAAGTTTCTCGGCTTTGCTAAATAATGTTATTCGAGGTGTTACTGTTCTGA 360
QY      361 TGTTCGAGGTTTGTGCTCAAGGTGGTTCACTGAGAGCGGCGGCTCAGAAATTTTC 420
Db      361 TGTTCGAGGTTTGTGCTCAAGGTGGTTCACTGAGAGCGGCGGCTCAGAAATTTTC 420
QY      421 TGAAGACAAGTTTGTGCTAAAGTTAAGATGGAGATCGTGGTTAGCAAGACCAAGTTGA 480
Db      421 TGAAGACAAGTTTGTGCTAAAGTTAAGATGGAGATCGTGGTTAGCAAGACCAAGTTGA 480
QY      481 GGATGTTATAGAAAATCATTGAGGAGGCAAGACTGGAGAGATTGGAGACGCGCAAGAT 540
Db      481 GGATGTTATAGAAAATCATTGAGGAGGCAAGACTGGAGAGATTGGAGACGCGCAAGAT 540
QY      541 TTTCTTGCTGCTGCTTTTTCAGATGTAATAAGAGTCCGCACTGGTGGAGCGGGGTGATAAGGC 600
Db      541 TTTCTTGCTGCTGCTTTTTCAGATGTAATAAGAGTCCGCACTGGTGGAGCGGGGTGATAAGGC 600
QY      601 TGAGAGGATGACAGGAGGCGGATCTGACATGAGTACTTCTGCTTGACTGCTGTGACCCAGC 660
Db      601 TGAGAGGATGACAGGAGGCGGATCTGACATGAGTACTTCTGCTTGACTGCTGTGACCCAGC 660
QY      661 AATATAGCATTCAGGACTAACTGTCCTTTGAGAAAGCCCGCCCTTATTAGCCATTATCC 720
Db      661 AATATAGCATTCAGGACTAACTGTCCTTTGAGAAAGCCCGCCCTTATTAGCCATTATCC 720
QY      721 AGTATAGCTTGATAATTTGAATTTTGTCTTAACTAAAGAAACAAAGATCTTTTCA 780
Db      721 AGTATAGCTTGATAATTTGAATTTTGTCTTAACTAAAGAAACAAAGATCTTTTCA 780
QY      781 TTATCCTGTTGATGATAATTTGAAAACGGAAGGATCGGCAATTTGTTCAAGTCTTGCAG 840
Db      781 TTATCCTGTTGATGATAATTTGAAAACGGAAGGATCGGCAATTTGTTCAAGTCTTGCAG 840
QY      841 ATAAATAACAAGAGAGGAGTAATGTTAA CAAAAAATAAAAAAATAAAAACTCGAG 897
Db      841 ATAAATAACAAGAGAGGAGTAATGTTAA CAAAAAATAAAAAAATAAAAACTCGAG 897
```

RESULT 2
AAF58584
ID AAF58584 standard; cDNA; 594 BP.
XX
AC AAF58584;
XX

```
DT      23-APR-2001 (first entry)
XX
DE      Ricinus communis P-P-II cDNA fragment.
XX
KW      Castor bean; P-II; plant nitrogen regulatory gene; P-P-II;
KW      nitrogen assimilation; transgenic plant; herbicide screening; ss.
XX
OS      Ricinus communis.
XX
PN      US6177275-B1.
XX
PD      23-JAN-2001.
XX
PF      23-JUL-1997; 97US-00899330.
PR      24-JUL-1996; 96US-0022328P.
XX
PA      (UYNV ) UNIV NEW YORK STATE.
PI      Coruzzi GM, Lam H, Hsieh M;
XX
XX      WEI; 2001-158572/16.
DR      P-PSDB; AAB69496.
XX
PT      Novel P-P-II genes capable of regulating plant nitrogen assimilation,
PT      useful for transgenic plant production, and as probes for isolating
PT      additional genomic clones having P-P-II gene promoters.
XX
PS      Claim 2; Col 37-38; 35pp; English.
XX
CC      The present sequence encodes a nitrogen regulatory P-II protein. Novel
CC      plant PII (also called P-P-II) nucleotide sequences have been isolated.
CC      They are useful for regulating nitrogen assimilation in plants, and in
CC      transgenic plant production. They are also used to engineer organisms
CC      that overexpress wild-type or mutant P-P-II regulatory proteins. P-P-II
CC      proteins are useful for in vitro screening of herbicides. P-P-II
CC      nucleotides may be used as probes for isolating additional genomic clones
CC      with the promoters of P-P-II genes. P-P-II promoters are light- and/or
CC      sucrose-inducible, and are suitable for genetic engineering of plants
XX
SQ      Sequence 594 BP; 149 A; 117 C; 157 G; 171 T; 0 U; 0 Other;

      Query Match      66.2%; Score 594; DB 4; Length 594;
      Best Local Similarity 100.0%; Pred. No. 5e-137;
      Matches 594; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      50 CGGCACGAGGCTACTGCGAAACTGGGCTTGCTCACTCTCTTCAATTAATCAATCAAG 109
Db      1 CGGCACGAGGCTACTGCGAAACTGGGCTTGCTCACTCTCTTCAATTAATCAATCAAG 60
QY      110 AAGAATTCCTGTTTGTGATTTCAAGCGCGTAAGATATGCCCGCTCGTTCTGATTAAT 169
Db      61 AAGAATTCCTGTTTGTGATTTCAAGTTTGTGTTGTCAGAGCTTAGACATTTCTCGGTT 120
QY      170 TCTCACTTTAACACCGCGGTCAAGCGCGTAAGATATGCCCGCTCGTTCTGATTAAT 229
Db      121 TCTCACTTTAACACCGCGGTCAAGCGCGTAAGATATGCCCGCTCGTTCTGATTAAT 180
QY      230 GCCCAAAGCTCGCCTGACTACATTTCTGATGCTAAATTTACAAAGTGAAGCAATTTCTC 289
Db      181 GCCCAAAGCTCGCCTGACTACATTTCTGATGCTAAATTTCTACAAAGTGAAGCAATTTCTC 240
QY      290 AGGCCCTGGCGAGTCTCGCAAGTTTCTCGGCTTTGCTAAATAATTTGTTATTCGAGGTT 349
Db      241 AGGCCCTGGCGAGTCTCGCAAGTTTCTCGGCTTTGCTAAATAATTTGTTATTCGAGGTT 300
QY      350 ACTGTTTCTGATGTTTCGAGGTTTGGTGTCTCAAGGTGTTCAACTGAGAGCGAGGCGGC 409
Db      301 ACTGTTTCTGATGTTTCGAGGTTTGGTGTCTCAAGGTGTTCAACTGAGAGCGAGGCGGC 360
QY      410 TCAGAAATTTTCTGAAGACAAGTTTGTGCTAAAGTTAAGATGAGATCGTGGTTAGCAA 469
Db      361 TCAGAAATTTTCTGAAGACAAGTTTGTGCTAAAGTTAAGATGAGATCGTGGTTAGCAA 420
```


CC transgenic plant production. They are also used to engineer organisms
CC that overexpress wild-type or mutant P-II regulatory proteins. P-II
CC proteins are useful for in vitro screening of herbicides. P-II
CC nucleotides may be used as probes for isolating additional genomic clones
CC with the promoters of P-II genes. P-II promoters are light- and/or
CC sucrose-inducible, and are suitable for genetic engineering of plants
XX
SQ Sequence 588 BP; 164 A; 105 C; 147 G; 172 T; 0 U; 0 Other;

Query Match 31.8%; Score 285; DB 4; Length 588;
Best Local Similarity 70.3%; Pred. No. 1.5e-60;
Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

QY 56 GAGGCTACTGCGAAACTGGGCTTGGCTCACTCTCTTCAATTAACATCAAGAAAGAA 115
DB 7 GCGTCAATGACGAAACCCATCTCAATAAATCTCTCGGTTTCTATCTGATCGAAAGAAC 66

QY 116 TTCCCTGTTTGTGATTTTGTGTTTGTGTCAGAGCTTAGACATTTCTCGGTTTCTCAC 175
DB 67 ATTGCTTTCTGATTCGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 126

QY 176 TTTAACACCGCGGTCAAGCGGTAAAGATATGCC---CCGTCGTTCTGTGATTAATGCC 232
DB 127 CTCGATTTGGTCACAAAGTCACCGAGTAAATACAGTGTGTTTACCTGTCGTAGTGC 186

QY 233 CAAAGCTCGCTGACTACATTCCTGATGCTAAATTTACAAAGTGAAGCAATTTCTCAGG 292
DB 187 CAAATATCTTCTGATTAATTTCCAGACTCGAAATTTTACAAGGTGGAAGCAATTTGT 246

QY 293 CCCTGGGAGTCTCGCAAGTTTCTCGGCTTTGCTAAATAATTTGTTATTCGAGGTGTTACT 352
DB 247 CCATGGAGATCCAGCAAGTTTTCATCGGCTTTACTGAAATCGGGATTCGAGGTGTTACT 306

QY 353 GTTCTGATGTTTCGAGGTTTGGTGTCTCAAGTGGTTCAACTGAGAGCGGCGGCTCA 412
DB 307 GTTCTGATGTTTCGAGGTTTGGTGTCTCAAGTGGTTCAACTGAGAGCGGCGGCTCT 366

QY 413 GAATTTTGAAGACAAGTTTGTGTTGCTAAAGTTAAGTGAAGTTCGTTAGCAAGAAC 472
DB 367 GAGTTCTCGAAGACAAGTTTGTGTTGCTAAAGTTAAGTGAAGTTCGTTAGCAAGAAC 426

QY 473 CAGGTTGAGGATGTTATAGAAAAATCAATTCAGGAGGCAAGAACTGGAGAGATTGGAGAC 532
DB 427 CAAAGTGAATCTGTAATCAACACAATAATTAAGAGGCAAGGACAGGAGAGATTGGTGT 486

QY 533 GGCAAGATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 592
DB 487 GGCAAGATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 546

QY 593 GATAAGGCTGAGAGGATGACAGGAG 617
DB 547 GAGAAAGCAGAGAGATGACTGGTG 571

RESULT 5
ABZ14328
ID ABZ14328 standard; DNA; 591 BP.
XX
AC ABZ14328;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2133.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX
PN W0200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US026685.

XX
PR 24-AUG-2000; 2000US-0227866P.
PR 26-JAN-2001; 2001US-0264647P.
PR 22-JUN-2001; 2001US-0300111P.
XX
PA (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX
DR WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed and
XX producing plants with increased tolerance to these abiotic stresses.
PS Claim 144; SEQ ID NO 2133; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 591 BP; 165 A; 105 C; 148 G; 173 T; 0 U; 0 Other;

Query Match 31.8%; Score 285; DB 6; Length 591;
Best Local Similarity 70.3%; Pred. No. 1.5e-60;
Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

QY 56 GAGGCTACTGCGAAACTGGGCTTGGCTCACTCTCTTCAATTAACATCAAGAAAGAA 115
DB 7 GCGTCAATGACGAAACCCATCTCAATAAATCTCTCGGTTTCTATCTGATCGAAAGAAC 66

QY 116 TTCCCTGTTTGTGATTTTGTGTTTGTGTCAGAGCTTAGACATTTCTCGGTTTCTCAC 175
DB 67 ATTGCTTTCTGATTCGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 126

QY 176 TTTAACACCGCGGTCAAGCGGTAAAGATATGCC---CCGTCGTTCTGTGATTAATGCC 232
DB 127 CTCGATTTGGTCACAAAGTCACCGAGTAAATACAGTGTGTTTACCTGTCGTAGTGC 186

QY 233 CAAAGCTCGCTGACTACATTCCTGATGCTAAATTTACAAAGTGAAGCAATTTCTCAGG 292
DB 187 CAAATATCTTCTGATTAATTTCCAGACTCGAAATTTTACAAGGTGGAAGCAATTTGT 246

QY 293 CCCTGGGAGTCTCGCAAGTTTCTCGGCTTTGCTAAATAATTTGTTATTCGAGGTGTTACT 352
DB 247 CCATGGAGATCCAGCAAGTTTGGTGTCTCAAGTGGTTCAACTGAGAGCGGCGGCTCT 306

QY 353 GTTCTGATGTTTCGAGGTTTGGTGTCTCAAGTGGTTCAACTGAGAGCGGCGGCTCA 412
DB 307 GTTCTGATGTTTCGAGGTTTGGTGTCTCAAGTGGTTCAACTGAGAGCGGCGGCTCT 366

QY 413 GAATTTTGAAGACAAGTTTGTGTTGCTAAAGTTAAGTGAAGTTCGTTAGCAAGAAC 472
DB 367 GAGTTCTCGAAGACAAGTTTGTGTTGCTAAAGTTAAGTGAAGTTCGTTAGCAAGAAC 426

QY 473 CAGGTTGAGGATGTTATAGAAAAATCAATTCAGGAGGCAAGAACTGGAGAGATTGGAGAC 532
DB 427 CAAAGTGAATCTGTAATCAACACAATAATTAAGAGGCAAGGACAGGAGAGATTGGTGT 486

QY 533 GGCAAGATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 592
DB 487 GGCAAGATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 546

QY 593 GATAAGGCTGAGAGGATGACAGGAG 617

Db 547 GAGAAAGCAGAGAAGATGACTGGTG 571

RESULT 6

AAV21209_01
Continuation (2 of 17) of AAV21209 from base 100001 (Methanococcus jannaschii circular
WP Sequence split into 17 fragments LOCUS AAV21209 Accession Aav21209
WP Fragment Name Begin End
WP AAV21209_00 1 110000
WP AAV21209_01 100001 210000
WP AAV21209_02 200001 310000
WP AAV21209_03 300001 410000
WP AAV21209_04 400001 510000
WP AAV21209_05 500001 610000
WP AAV21209_06 600001 710000
WP AAV21209_07 700001 810000
WP AAV21209_08 800001 910000
WP AAV21209_09 900001 1010000
WP AAV21209_10 1000001 1110000
WP AAV21209_11 1100001 1210000
WP AAV21209_12 1200001 1310000
WP AAV21209_13 1300001 1410000
WP AAV21209_14 1400001 1510000
WP AAV21209_15 1500001 1610000
WP AAV21209_16 1600001 1664976

Query Match 10.7%; Score 96.4; DB 2; Length 110000;
Best Local Similarity 57.9%; Pred. No. 4.4e-13;
Matches 191; Conservative 0; Mismatches 136; Indels 3; Gaps 1;

QY 263 AAATTCTACAAAGTGAAGCAATTTCTCGAGCTTCTGATGTTTCGAGGTTTTCGCTCAA 382
Db 80891 AAAATGAAAAAGTTGAAGCAATCATAGACCGGAGAGTTGGAGTTGTTAAAAAGGCT 80950

QY 323 TTGCTAAAAAATTCGATTCGAGGTTTCTGATGTTTCGAGGTTTTCGCTCAA 382

Db 80951 TTGCTGATGCTGGATATGTTGGAATGACTGTTAGTAGAGGTTAAGGTTAGGGAGTTCAA 81010

QY 383 GGTGTTCAACTGAGAGCGGCGGCTCAGAAATTTCTGAAGACAAGTTTGTGCTAAA 442

Db 81011 GGTGGAATAGTTGAGAGGTATAGGGGAGAGATATGTTGTTG---ATTAAATCCAAAG 81067

QY 443 GTTAAGATGGAGATCGTGTAGCAAGACCGGTTGAGGATGTTATAGAAAAATCATT 502

Db 81068 GTTAAGATTGAATTGGTTGTTAAAGAGGAGAGATGTTGATAAGTTATGATATATATGC 81127

QY 503 GAGGAGGCAAGAACTGGAGAGATTGGAGACGGCAAGATTTCTGCTGCTGTTTCAGAT 562

Db 81128 GAGAATGCAAGAACAGAAACCCAGGAGATGGAAAAATCTTCGTATACCAAGTAGAAGA 81187

QY 563 GTAATAAGAGTCCGCACTGCTGAGCGGGGT 592

Db 81188 GTCGTAAGAGTAAAGAACAAAAAGAGAGGTT 81217

RESULT 7

AAV21209_14/c
Continuation (15 of 17) of AAV21209 from base 1400001 (Methanococcus jannaschii circular
WP Sequence split into 17 fragments LOCUS AAV21209 Accession Aav21209
WP Fragment Name Begin End
WP AAV21209_00 1 110000
WP AAV21209_01 100001 210000
WP AAV21209_02 200001 310000
WP AAV21209_03 300001 410000
WP AAV21209_04 400001 510000
WP AAV21209_05 500001 610000
WP AAV21209_06 600001 710000
WP AAV21209_07 700001 810000
WP AAV21209_08 800001 910000
WP AAV21209_09 900001 1010000
WP AAV21209_10 1000001 1110000
WP AAV21209_11 1100001 1210000

WP AAV21209_12 1200001 1310000
WP AAV21209_13 1300001 1410000
WP AAV21209_14 1400001 1510000
WP AAV21209_15 1500001 1610000
WP AAV21209_16 1600001 1664976

Query Match 10.7%; Score 96.2; DB 2; Length 110000;
Best Local Similarity 56.3%; Pred. No. 4.9e-13;
Matches 201; Conservative 0; Mismatches 153; Indels 3; Gaps 1;

QY 244 TGACTACATTCCTGATGCTAAATTTCTACAAAGTGAAGCAATTTCTCAGGCGCTGGCAGT 303

Db 13404 TTATAAAATGAGGTTGAGATTATGAAAAAGTTGAAGCAATCATAGACCGGAGAGTT 13345

QY 304 CTCGCAAGTTTCTCGGCTTTGCTAAATAATGGTATTGAGGTTGTTACTGTTCTGATGT 363

Db 13344 GGAGATTGTTAAAAAGGCTTTGCTCTGATGCTGGGTATGTTGAATGACTGTTAGTGAGT 13285

QY 364 TCGAGGTTTTCGCTCAAGGTGGTTCAACTGAGAGGCGGCGCTCAGAAATTTCTGA 423

Db 13284 TAAGGTTAGGGAGTTCAAGGTGGAATAGTTGAGAGGTATAGGGGAGAGATATATTGT 13225

QY 424 AGACAAGTTTGTGCTAAAGTTAAGATGGAGATCGTGGTTAGCAAAAGACCGGTTGAGGA 483

Db 13224 TG---ATTAAATCCAAAGGTTAAGATTGAGTTGGTTGTTAAAGAGGAAGATGTTGATTA 13168

QY 484 TGTTATAGAAAAATCATTCAGGAGCGCAAGAACTGGAGATTCGAGACCGGCAAGATTTT 543

Db 13167 TGTTATTGATATCATATGCGAGAATGCAAGAACAGGAACCCAGGAGATGGAAAAATCTT 13108

QY 544 CTGCTGCTGTTTCAGATGTAATAAGAGTCCGCACTGTTGAGCGGGGTGATAAGGC 600

Db 13107 CGTCATACCAAGTAGAAAGAGTCTGTAAGAGTAAAGAACAAAGAAAGAAAGAGGC 13051

RESULT 8

AAF28548

ID AAF28548 standard; DNA; 96109 BP.

XX AAF28548;

DT 04-APR-2001 (first entry)

XX Genomic fragment #35.

DE Genomic library; bacteria; human upper airway; otitis media; sinusitis;
XX bronchopulmonary; endocarditis; meningitis; ss.

OS Moraxella catarrhalis.

XX WO200078968-A2.

XX 28-DEC-2000.

XX 16-JUN-2000; 2000WO-USC16649.

XX 18-JUN-1999; 99US-0140121P.

XX (INCY-) INCYTE GENOMICS INC.

XX Lagace RE, Patterson C, Berg KL;

XX WPI; 2001-041427/05.

XX Genomic library for identifying diagnostic and therapeutic compositions,
XX and for identifying virulence factors, regulatory elements and drug
XX targets, comprises Moraxella catarrhalis nucleic acids.

PS Claim 1; Page 345-368; 545pp; English.

XX The present invention relates to a Moraxella catarrhalis genomic library
XX comprising of a combination of 41 nucleic acid molecules (see AAF28514-
XX AAF28554). The library has a number of uses described in the

PI Buchrieser C;
XX WPI; 2003-148459/14.
DR Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
XX useful e.g. as therapeutic antimicrobials and agricultural pesticides.
PS Claim 2; SEQ ID NO :0494; 1205pp; French.
XX
CC The invention relates to the isolation of genes and their encoded
CC proteins from Photorhabdus luminescens. The isolated sequences are
CC sources of probes and primers for detecting the genome of P. luminescens
CC and related species; to study polymorphisms; for gene analysis and for
CC detection/amplification of the genes. Antibodies (Ab) raised against the
CC polypeptides encoded by the genes are used for detection/identification
CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
CC carry a gene-containing vector are used to select compounds that
CC modulate, regulate, induce or inhibit expression of the genes in plants,
CC animals or microorganisms other than P. luminescens and are able to alter
CC response or sensitivity to toxins and antibiotics produced by P.
CC luminescens. Cells transformed to express the genes are useful for
CC recombinant production of the proteins, particularly toxins and
CC antibacterials useful as insecticides, bactericides and fungicides. The
CC genes, proteins, vectors containing the genes and Ab are also useful
CC therapeutically (to treat microbial infection by bacteria or fungi that
CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC biopesticides. Other uses of the genes and the proteins are as virulence
CC factors and for identifying targets of human diseases for which P.
CC luminescens is a model (particularly plague and whooping cough). This
CC sequence represents one of the isolated P. luminescens genes
XX
SQ Sequence 339 BP; 105 A; 51 C; 96 G; 87 T; 0 U; 0 Other;

Query Match 9.1%; Score 81.4; DB 7; Length 339;
Best Local Similarity 55.0%; Pred. No. 3.1e-10;
Matches 182; Conservative 0; Mismatches 146; Indels 3; Gaps 1;

QY 270 ACAAGTGAAGCAATCTCAGGCCCTGGCGAGTCTCGCAAGTTTCCTCGGCTTTGCTAA 329
Db 5 AAAGATTGATGCGATTATCAAACTTTCAAAATTAGATGATGCGTGAAGCTCTGGCGG 64

QY 330 AAATTGGTATTCGAGGTGTTACTGTTTCTGATGTTTCGAGGTTTGGTGTCAAGGTGGTT 389
Db 65 AAGTGGGTATCACCGGAATGACAGTAAACAGAGGTGAAGGTTTGGCGCCAAAGGTC 124

QY 390 CAAGTGAAGGCGGCGCTCAGAAATTTCTGAAGACAAAGTTTGTGCTAAAGTTAAGA 449
Db 125 ATACAGAGCTGTATCGCGGTGCAATATATGTTGGAT---TTCTGCCAAAGTGAAGA 181

QY 450 TGGAGATCGTGGTTAGCAAGACCCAGGTTGAGGATGTTATAGAAAATCATTTGAGGAGG 509
Db 182 TAGAAATTGTCGTGCCAGATGATATTGTCGATACCTGTGTTGAACCATTTATGCAGCGG 241

QY 510 CAAGAACTGGAGAGATTGGAGACGGCAAGATTTTCTGCTGCCTGTTTCAGATGTAATAA 569
Db 242 CACAGACCGGGAATCGGTGATGGTAAATATTTGTAATTGATGTAGCACGTTGTGTC 301

QY 570 GAGTCCGCACTGGTGACCGGGGTGATAAGGC 600
Db 302 GTATCCGACCGGTGAGCAGGATGAAGAGGC 332

RESULT 11
ACF67367_52

Continuation (53 of 57) of ACF67367 from base 5200001 (Photorhabdus luminescens nucleoti
WP Sequence split into 57 fragments LOCUS ACF67367 Accession ACF67367
WP Fragment Name Begin End
WP ACF67367_00 1 110000
WP ACF67367_01 100001 210000
WP ACF67367_02 200001 310000
WP ACF67367_03 300001 410000
WP ACF67367_04 400001 510000
WP ACF67367_05 500001 610000

WP ACF67367_06 600001 710000
WP ACF67367_07 700001 810000
WP ACF67367_08 800001 910000
WP ACF67367_09 900001 1010000
WP ACF67367_10 1000001 1110000
WP ACF67367_11 1100001 1210000
WP ACF67367_12 1200001 1310000
WP ACF67367_13 1300001 1410000
WP ACF67367_14 1400001 1510000
WP ACF67367_15 1500001 1610000
WP ACF67367_16 1600001 1710000
WP ACF67367_17 1700001 1810000
WP ACF67367_18 1800001 1910000
WP ACF67367_19 1900001 2010000
WP ACF67367_20 2000001 2110000
WP ACF67367_21 2100001 2210000
WP ACF67367_22 2200001 2310000
WP ACF67367_23 2300001 2410000
WP ACF67367_24 2400001 2510000
WP ACF67367_25 2500001 2610000
WP ACF67367_26 2600001 2710000
WP ACF67367_27 2700001 2810000
WP ACF67367_28 2800001 2910000
WP ACF67367_29 2900001 3010000
WP ACF67367_30 3000001 3110000
WP ACF67367_31 3100001 3210000
WP ACF67367_32 3200001 3310000
WP ACF67367_33 3300001 3410000
WP ACF67367_34 3400001 3510000
WP ACF67367_35 3500001 3610000
WP ACF67367_36 3600001 3710000
WP ACF67367_37 3700001 3810000
WP ACF67367_38 3800001 3910000
WP ACF67367_39 3900001 4010000
WP ACF67367_40 4000001 4110000
WP ACF67367_41 4100001 4210000
WP ACF67367_42 4200001 4310000
WP ACF67367_43 4300001 4410000
WP ACF67367_44 4400001 4510000
WP ACF67367_45 4500001 4610000
WP ACF67367_46 4600001 4710000
WP ACF67367_47 4700001 4810000
WP ACF67367_48 4800001 4910000
WP ACF67367_49 4900001 5010000
WP ACF67367_50 5000001 5110000
WP ACF67367_51 5100001 5210000
WP ACF67367_52 5200001 5310000
WP ACF67367_53 5300001 5410000
WP ACF67367_54 5400001 5510000
WP ACF67367_55 5500001 5610000
WP ACF67367_56 5600001 5648894

Query Match 9.1%; Score 81.4; DB 7; Length 110000;
Best Local Similarity 55.0%; Pred. No. 2.3e-09;
Matches 182; Conservative 0; Mismatches 146; Indels 3; Gaps 1;

QY 270 ACAAGTGAAGCAATCTCAGGCCCTGGCGAGTCTCGCAAGTTTCCTCGGCTTTGCTAA 329
Db 49307 AAAGATTGATGCGATTATCAAACTTTCAAAATTAGATGATGCGTGAAGCTCTGGCGG 49366

QY 330 AAATTGGTATTCGAGGTGTTACTGTTTCTGATGTTTCGAGGTTTGGTGTCTCAAGGTGGTT 389
Db 49367 AAGTGGGTATCACCGGAATGACAGTAAACAGAGGTGAAGGTTTGGCGGCCCAAAAGGTC 49426

QY 390 CAAGTGAAGGCGGCGGCTCAGAAATTTCTGAAGACAAAGTTTGTGCTAAAGTTAAGA 449
Db 49427 ATACAGAGCTGTATCGCGGTGCAATATATGTTGGAT---TTCTGCCAAAGTGAAGA 49483

QY 450 TGGAGATCGTGGTTAGCAAGACCCAGGTTGAGGATGTTATAGAAAATCATTTGAGGAGG 509
Db 49484 TAGAAATTGTCGTGCCAGATGATATTGTCGATACCTGTGTTGAACCATTTATGCAGACGG 49543

QY 510 CAAGAACTGGAGAGATTGGAGACGGCAAGATTTTCTGCTGCCTGTTTCAGATGTAATAA 569

Db 49544 CACAGACCGGAAATCGGTGATGGTAAATATTCTGATTGATGACGCGTGTGTC 49603

QY 570 GAGTCCGCACTGGTACGCGGGGTGATAAGGC 600

Db 49604 GTATCCGCAACCGGTGACGAGGATGAAGAGGC 49634

RESULT 12

ACF65387_3

Continuation (4 of 7) of ACF65387 from base 300001 (Photorhabdus luminescens nucleotide

WP Sequence split into 7 fragments LOCUS ACF65387 Accession ACF65387

WP Fragment Name Begin End

WP ACF65387_0 1 110000

WP ACF65387_1 100001 210000

WP ACF65387_2 200001 310000

WP ACF65387_3 300001 410000

WP ACF65387_4 400001 510000

WP ACF65387_5 500001 610000

WP ACF65387_6 600001 696798

Query Match 9.1%; Score 81.4; DB 7; Length 110000;

Best Local Similarity 55.0%; Pred. No. 2.3e-09;

Matches 182; Conservative 0; Mismatches 146; Indels 3; Gaps 1;

QY 270 ACAAAGTGAAGCAATCTCAGGCCCTGGCGAGTCTCGCAAGTTTCTCGGCTTTGCTAA 329

Db 88821 AAAAGATTGATCGGATTATCAAACTTTCAATTAGATGATGCGTGAAGCTCTGGCGG 88880

QY 330 AAATTGGTATTCGAGGTGTTACTGTTCTGATGTTTCGAGGTTTTCGCTCAAGTGGTT 389

Db 88881 AASTGGGTATCACCGGAATGACAGTAACAGAGGTGTAAGGTTTGGCGGCCAAAAGGTC 88940

QY 390 CAACTGAGAGCGGCGGCTCAGAAATTTCTGAAGACAAAGTTTGTGCTAAAGTTAAGA 449

Db 88941 ATACAGAGCTGTATCGCGGTGCAGATATATGTTGGAT---TTTCTGCCAAAAGTGAAGA 88997

QY 450 TGGAGATCGTGGTTAGCAAGACCAAGTTTCTGAGGATGTTGAGGATGTTATAGAAAAATCATTCAGGAGG 509

Db 88998 TAGAAATTGTGTCGCCAGATGATATTGTCGATACCTGTGTTGAAACCATTTATGCAGACGG 89057

QY 510 CAAGAAGTGGAGAGATTGGAGACGGCAAGATTTTCTGCTGCTGCTTTTCAGATGTAATAA 569

Db 89058 CACAGACCGGAAATCGGTGATGTTAAATATTTCGATTGATGATGACGCGTGTGTC 89117

QY 570 GAGTCCGCACTGGTACGCGGGGTGATAAGGC 600

Db 89118 GTATCCGCAACCGGTGACGAGGATGAAGAGGC 89148

WP AAT42063_18 1800001 1830121

Query Match 8.7%; Score 77.6; DB 2; Length 110000;

Best Local Similarity 55.1%; Pred. No. 2e-08;

Matches 174; Conservative 0; Mismatches 139; Indels 3; Gaps 1;

QY 270 ACAAAGTGAAGCAATTTCTAGGCCCTGGCGAGTCTCGCAAGTTTCTCGGCTTTGCTAA 329

Db 63919 AAAAATCGAAGCAATGATTAACCCCTTTAAATTAGACGATGTCGAGAAAGTCTTTTCAG 63978

QY 330 AAATTGGTATTCGAGGTGTTACTGTTTCTGATGTTTCGAGGTTTTCGCTCAAGTGGTT 389

Db 63979 ATATTGGTATTTTCAGGTATGACAATCACAGAAAGTACGCGGATTTGGTCGTCAAAAGGTC 64038

QY 390 CAACTGAGAGCGGCGGCTCAGAAATTTCTGAAGACAAAGTTTGTGCTAAAGTTAAGA 449

Db 64039 ATACAGAACTTTATCGTGTGTCGGAATATATGTTGGAT---TTTCTGCCGAAAGTGAAT 64095

QY 450 TGGAGATCGTGGTTAGCAAGACCAAGTTTTCGAGGATGTTGAGGATGTTATAGAAAAATCATTCAGGAGG 509

Db 64096 TGAAGTGGTAGTTCTCTGATGAGCTTGTGGATCAATGATGATGAAGCGATTATTGAAACGG 64155

QY 510 CAAGAAGTGGAGAGATTGGAGACGGCAAGATTTTCTGCTGCTGCTTTTCAGATGTAATAA 569

Db 64156 CACAAACAGGTAAATCGGTGACGGCAAAATTTTGTATTATCAGTTGAGAGAGCCATCC 64215

QY 570 GAGTCCGCACTGGTGA 585

Db 64216 GCATTGCAACGGCGGA 64231

RESULT 14

ABZ40022

ID ABZ40022 standard; DNA; 363 BP.

XX AC ABZ40022;

XX AC ABZ40022;

DT 07-MAR-2003 (first entry)

XX DE N. gonorrhoeae nucleotide sequence SEQ ID 4633.

XX DE N. gonorrhoeae nucleotide sequence SEQ ID 4633.

XX KW Antibacterial; infection; vaccine; gene therapy; gene; ds.

XX OS Neisseria gonorrhoeae.

XX OS Neisseria gonorrhoeae.

XX PN WO200279243-A2.

XX PD 10-OCT-2002.

XX PF 12-FEB-2002; 2002WO-IB002069.

XX PR 12-FEB-2001; 2001GB-000003424.

XX PA (CHIR-) CHIRON SPA.

XX PI Fontana MR, Pizza M, Maignani V, Monaci E;

XX DR WPI; 2003-058415/05.

XX DR P-PSDB; ABP79052.

XX PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a

XX PT medicament for treating or preventing N. gonorrhoeae infection.

XX PS Disclosure; Page 528; 815pp; English.

XX CC The present invention relates to proteins from Neisseria gonorrhoeae.

XX CC Also disclosed are the nucleic acid molecules encoding the proteins and

XX CC antibodies that specifically bind to the proteins. The composition

XX CC comprising the protein, nucleic acid or antibody is useful for the

XX CC manufacture of a medicament for treating or preventing N. gonorrhoeae

XX CC infection, this may be in the form of a vaccine or gene therapy.

XX CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid

XX CC molecules of the invention

```
XX Sequence 363 BP; 84 A; 77 C; 123 G; 79 T; 0 U; 0 Other;
SQ
Query Match      8.6%; Score 77; DB 7; Length 363;
Best Local Similarity 54.1%; Pred. No. 3.9e-09;
Matches 180; Conservative 0; Mismatches 150; Indels 3; Gaps 1;

QY 270 ACAAGTGGAGCAATTCAGGCCCTCGGAGTCTCGCAAGTTTCCTCGGCTTTGCTAA 329
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 32 AAAAAATCGAGCGATTGTCAAACCGTTCAGCTCGACGCTGCGCGAGCGGTGACGG 91

QY 330 AAATTGGTATTCGAGGTGTACTGTTCTGATGTTTCGAGGTTTTCGCTCAAGGTGTT 389
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 92 AAATCGGCATTACGGGCATGACCGTCAAGGAGGTCARAGGTTTCGGCAGGCAAGGGGC 151

QY 390 CAACTGAGAGGAGGGCGGCTCAGAAATTTCTCAAGACAAAGTTTGTGCTAAAGTTAAGA 449
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152 ATACGGAATCTATCGGGTGCAGTACGCGGTGAT---TTCCTGCCCAAGGTCAAAA 208

QY 450 TGGAGATCGTGTAGCAAGACCAGGTTGAGGATGTTATAGAAAAATCAATTGAGGAGG 509
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 209 TCGAGTTGGTGTGGCGGATGATGCCGTGCAAGCGCGGATTCACGTGATTTTCGAGGTGG 268

QY 510 CAAGAACTGGAGATGAGAGCGGCAAGATTTCTTGCTGCCTGTTTCAGATGTAATAA 569
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 269 CGCGTTCGGGCAAAATCGGCAAGCGCAAGATTTTGTGCTGCCGTGCGGAGGCAATCC 328

QY 570 GAGTCGGCACTGTGAGCGGGGTGATAAGGCTG 602
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 329 GTATCCGACCGGCGCAACGTTCCGACGCGGCGG 361

RESULT 15
AAA81472
ID AAA81472 standard; DNA; 16526 BP.
XX
AC AAA81472;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_20 SEQ ID NO:20.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX
EF 08-OCT-1999; 99WO-US023573.
XX
PR 09-OCT-1998; 98US-0103794P.
XX
PR 30-APR-1999; 99US-0132068P.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Maignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
DR WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be used
PT in the diagnosis and treatment of N. meningitidis infection and other
PT Neisserial infections, for example, N.gonorrhoea.
XX
PS Claim 7; Page 466-471; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic proteins
CC from Neisseria genomic sequences. AAA81453 to AAA82414 represent
```

```
CC specifically claimed Neisseria meningitidis genomic DNA sequences;
CC AAA81260 to AAA81303 and AAB25620 to AAB25663 represent Neisseria DNA
CC sequences and their corresponding proteins; AAA81254 to AAA81259 and
CC AAA81304 to AAA81321 represent PCR primers used in the isolation of
CC Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent
CC Neisseria meningitidis MenB polynucleotide ORF sequences, which are all
CC used in the exemplification of the present invention. The nucleic acid
CC sequences, protein sequences, and antibodies against them, can be used in
CC the manufacture of a composition. The composition can be used as a
CC medicament (or in the manufacture of a medicament) for treating,
CC preventing or diagnosing infection due to Neisserial bacteria. For
CC example, some of the identified proteins could be components of vaccines
CC against Meningococcus B; against all serotypes; and/or against all
CC pathogenic Neisseriae. Identification of sequences from the bacterium
CC will also facilitate production of biological probes, particularly
CC organism-specific probes. Attempts to make efficacious Meningococcus B
CC vaccines have failed mainly due to antigen tolerance. Multivalent
CC vaccines have also been tried but none have successfully overcome
CC antigenic variability. The provision of further, complete sequences may
CC provide an opportunity to identify secreted or surface exposed proteins
CC that may be presumed targets for the immune system and which are not
CC antigenically variable or at least more conserved than other more
CC variable regions
XX
SQ Sequence 16526 BP; 3183 A; 4006 C; 5067 G; 4268 T; 0 U; 2 Other;

Query Match      8.1%; Score 73; DB 3; Length 16526;
Best Local Similarity 55.3%; Pred. No. 1.4e-07;
Matches 184; Conservative 0; Mismatches 145; Indels 4; Gaps 2;

QY 270 ACAAAAGTGAAGCAATTCAGGCCCTCGGAGTCTCGAAGTTTCCTCGGCTTTGCTAA 329
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13603 AAAAAATCGAGCGATTGTCAAACCGTTCAAACCTCGACGAGCTGCGAGGCGTTGACGG 13662

QY 330 AAATTGGTATTCGAGGTGTACTGTTTCTGAAGTTCGAGGTTTGGTCTCAAGGTGTT 389
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13663 AAATCGGCATTACGGGCGATGACCGTCAGCGAGGTCAAAGGTTTCGGCAGGCAAGGGGC 13722

QY 390 CAACTGAGAGGCGCGGCTCAGAAATTTCTGAAGACAAAGTTTGTGCTAAAGTTAAGA 449
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13723 ATACGGAATCTATCGCGGCGCGAATACGCGGTGAT---TTCCTGCCCAAAATCAAAA 13779

QY 450 TGGAGATCGTGTGTTAGCAAGACCAGGTTGAGGATGTTATAGAAAAATCATTGAGGAGG 509
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13780 TCGAGCTGTTGTTGGCGGATGATGCTGTGGAACGCGCGATTGACGTGATT-TCGAGGTGG 13838

QY 510 CAAGAACTGGAGAGATTGAGACGGCAAGATTTTCTGCTGCTGTTTCAGATGTAATAA 569
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13839 CGCGTTCGGGAAAAATCGGCGACGGCAAGATTTTGTGCTGCGGTTGAGGAGGCAATCC 13898

QY 570 GAGTCCGCACCTGGTGAGCGGGGTGATAAGGCTG 602
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13899 GTATCCGCACGGGCGAACGTTCCGACGCGGCAG 13931
```

Search completed: May 27, 2004, 19:59:40
Job time : 315.835 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 19:07:55 ; Search time 2078.03 Seconds
(without alignments)
12890.278 Million cell updates/sec

Title: US-09-756-541-14
Perfect score: 897
Sequence: 1 GGGGTGCGCGCTCTAGA.....AAAAAATACTCGAG 897

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:
2: em_esthum:
3: em_estin:
4: em_estmu:
5: em_estov:
6: em_estpl:
7: em_estro:
8: em_htc:
9: gb_est1:
10: gb_est2:
11: gb_htc:
12: gb_est3:
13: gb_est4:
14: gb_est5:
15: em_estfun:
16: em_estom:
17: em_gss_hum:
18: em_gss_inv:
19: em_gss_pln:
20: em_gss_vrt:
21: em_gss_fun:
22: em_gss_mam:
23: em_gss_mus:
24: em_gss_pro:
25: em_gss_rod:
26: em_gss_phg:
27: em_gss_vrl:
28: gb_gss1:
29: gb_gss2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	409.8	45.6	542	14	T15252
2	321.4	35.8	782	14	CB289081
3	300.8	33.5	723	13	BQ798001
4	298.6	33.3	597	10	BE239550

5	298.6	33.3	714	12	BI311072	BI311072	EST531282
6	297	33.1	812	14	CA922829	CA922829	EST640547
7	293	32.7	612	10	BF645522	BF645522	NF036F08E
8	284.2	31.7	759	14	CF920467	CF920467	gmthRww3-
9	283	31.5	866	13	C81819	C81819	Citr
10	277.8	31.0	677	10	BE823431	BE823431	GM700019B
11	276	30.8	799	12	BM113473	BM113473	EST561009
12	275.8	30.7	711	12	BI932336	BI932336	EST552225
13	275	30.7	732	9	AJ558383	AJ558383	AJ558383
14	274.2	30.6	613	9	AW035791	AW035791	EST281945
15	274.2	30.6	648	9	AI773079	AI773079	EST254179
16	274.2	30.6	702	12	BI932123	BI932123	EST552012
17	274.2	30.6	741	10	AW738071	AW738071	EST339498
18	274.2	30.6	804	12	BI930336	BI930336	EST550225
19	274.2	30.6	810	12	BI930060	BI930060	EST549949
20	273.8	30.5	756	12	BG595705	BG595705	EST494383
21	273.6	30.5	736	12	BI932913	BI932913	EST552802
22	273.2	30.5	685	12	BI929601	BI929601	EST549490
23	272.6	30.4	630	10	AW929534	AW929534	EST338322
24	267	29.8	719	12	BJ573686	BJ573686	BJ573686
25	266.2	29.7	732	14	CD820903	CD820903	BN25.040D
26	265.4	29.6	657	14	CB349285	CB349285	CAB2SG000
27	265.4	29.6	701	14	CB349370	CB349370	CAB2SG000
28	264.6	29.5	458	10	BE210153	BE210153	so39d03.Y
29	264.6	29.5	669	14	CD822497	CD822497	BN25.045G
30	263	29.3	685	14	CD825321	CD825321	BN25.060G
31	263	29.3	745	14	CD819046	CD819046	BN20.047L
32	262.2	29.2	675	14	CB346213	CB346213	CAB2SG000
33	249.2	27.8	549	14	CB261070	CB261070	33-E9570-
34	248.2	27.7	560	14	CA514765	CA514765	KSC9030G0
35	246.4	27.5	579	10	AW223643	AW223643	EST300454
36	244.4	27.2	673	14	CB085731	CB085731	hg15b09.9
37	239.2	26.7	540	14	CB349912	CB349912	CAB2SG000
38	238.6	26.6	551	14	CB349990	CB349990	CAB2SG000
39	238.6	26.6	598	14	CB350294	CB350294	CAB2SG000
40	236.4	26.4	540	10	AW624562	AW624562	EST322507
41	233.4	26.0	532	14	CB350222	CB350222	CAB2SG000
42	227.6	25.4	564	13	BQ118918	BQ118918	EST604494
43	227.4	25.4	729	14	CF475242	CF475242	RTWW2.14
44	227.2	25.3	796	14	CF395143	CF395143	RTDS2.9.G
45	223.6	24.9	690	14	CA290108	CA290108	SCAGFL801

ALIGNMENTS

RESULT 1
T15252
LOCUS
DEFINITION
T15252 crs852 lambdaZAPST Ricinus communis cDNA clone pcrs852 similar to
nitrogen-regulatory protein, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
T15252.1 GI:14190796
Ricinus communis (castor bean)
Ricinus communis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Acalyphoideae;
Acalyphaeae; Ricinus.
1 (bases 1 to 542)
vandeLoe,F.J., Turner,S. and Somerville,C.
Expressed sequence tags from developing castor seeds
Plant Physiol. 108, 1141-1150 (1995)
Contact: Somerville CR
Carnegie Institution
Carnegie Institution, 290 Panama St, Stanford, CA 94305
Tel: 4153251521
Email: crs@andrew.stanford.edu
Seq primer: T3.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
vandeLoe,F.J., Turner,S. and Somerville,C.
Expressed sequence tags from developing castor seeds
Plant Physiol. 108, 1141-1150 (1995)
Contact: Somerville CR
Carnegie Institution
Carnegie Institution, 290 Panama St, Stanford, CA 94305
Tel: 4153251521
Email: crs@andrew.stanford.edu
Seq primer: T3.

FEATURES
source
Location/Qualifiers
1..542
/organism="Ricinus communis"

/mol_type="mRNA"
/strain="Baker 296"
/db_xref="taxon:3988"
/clone="pcrs852"
/clone_lib="lambdaZAPST"
/note="Vector: lambdaZAPII; Site_1: EcoRI; Site_2: XhoI;
Poly(A)+ RNA was purified from developing stage III to
stage V (Greenwood & Bewley, Can. J. Bot. 60:1751-1760,
1982) endosperm plus embryo of immature castor fruits.
cDNA was synthesized and cloned into lambdaZAPII according
to the instructions of the manufacturer (Stratagene).
synthesis was primed from the poly(A) tail, and cloned
directionally into XhoI (3') and EcoRI (5') sites. In few
cases, sequence data indicated that this directionality
was reversed. Partial cDNA clones predominate."

ORIGIN

Query Match 45.6%; Score 408.8; DB 14; Length 542;
Best Local Similarity 95.5%; Pred. No. 1.7e-54;
Matches 428; Conservative 0; Mismatches 18; Indels 2; Gaps 1;
QY 59 GCTACTGCGAACTGGCTTGTCTCACTCCTCTTCAATCTAATAACATCAAGAAAGATTTC 118
Db 1 GCTACTGCGAACTGGCTTGTCTCACTCCTCTTCAATCTAATAACATCAAGAAAGATTTC 60
QY 119 CCTGTTTTGATTTAGTTTGTGTTTGTCCAGAGCTTAGACATTTCTCGGTTTCTCATT 178
Db 61 CCTGTTTTTNAATTCAGTTTGTGTTTGTCCAGAGCTTAGACATTTCTCGGTTTCTCATT 120
QY 179 AACACCGCGGTCAAGCGCGTAAGATATGCCCGCTGCTTCTGTGATTATGCCCAAGC 238
Db 121 AACACCGCGGTCAAGCGCGTAAGATATGCCCGCTGCTTCTGTGATTATGCCCAAGC 180
QY 239 TCGCCTGACTACATTCCTGATGCTAAATTTCTACAAAGTGAAGCAATTTCTCAGGCCCTGG 298
Db 181 TCGCCTGACTACATTCCTGATGCTAAATTTCTACAAAGTGAAGCAATTTCTCAGGCCCTGG 240
QY 299 CGAGTCTCGCAAGTTTCTCGGCTTTGCTAAATAATTTGGTATTCGAGGTGTACTGTTCT 358
Db 241 CGAGTCTCGCAAGTTTCTCGGCTTTGCTAAATAATTTGGTATTCGAGGTGTACTGTTCT 300
QY 359 GATGTTGAGGTTTTTGTCTCAAGGTGTTCAACTGAGAGGCGGGCGGCTCAGAATTT 418
Db 301 GATGTTGAGGTTTTTGTCTCAAGGTGTTCAACTGAGAGGCGGGCGGCTCAGAATTT 360
QY 419 TCTGAAGACAAGTTTGTCTAAAGTTAAGTGAAGATCGTGGTTAGCAAGACCAAGTT 478
Db 361 NCTGAAGNCAAGTTTGTCTAAAGTAAAGTGAAGACCGGTGGT--NGCAAAGCCAGGTT 418
QY 479 GAGGATGTTATAGAAAAAATCATTGAGG 506
Db 419 GAGGNTGTTATTGGAAAAATNATTGGGG 446

RESULT 2
CB289081
LOCUS
DEFINITION V-B-112F06 VAN-Baker-1 Vitis aestivalis cDNA clone V-B-112F06 5',
mRNA sequence.
CB289081
VERSION CB289081.1 GI:28602822
KEYWORDS EST.
SOURCE Vitis aestivalis
ORGANISM Vitis aestivalis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.
REFERENCE 1 (bases 1 to 782)
AUTHORS Hou,H.S., Phanikanth,T.V., Kovacs,L. and Qiu,W.P.
TITLE Expressed sequence tags of young leaf tissues of a
disease-resistant Vitis aestivalis var. Norton
JOURNAL Unpublished (2003)
COMMENT Contact: Wenping Qiu

Department of Fruit Science
Southwest Missouri State University-Mountain Grove
9740 Red Spring Road, Mountain Grove, MO 65711-2999, USA
Tel: 417 926 4105
Fax: 417 926 6646
Email: weq070f@smsu.edu
Insert Length: 782 Std Error: 0.00
Plate: VAN-Baker-1-12 row: F column: 06
Seq primer: T3 PRIMER
High quality sequence stop: 782
POLYA=No.

FEATURES

source

Location/Qualifiers
1..782
/organism="Vitis aestivalis"
/mol_type="mRNA"
/cultivar="Norton"
/db_xref="taxon:3605"
/clone="V-B-112F06"
/tissue_type="Leaf"
/dev_stage="Young leaf"
/lab_host="XL10-Gold E.coli"
/clone_lib="VAN-Baker-1"
/note="Vector: pBluescript II SK (+) Phagemid; Site 1: Xho
I; Site 2: EcoR I; VAN-Baker-1 is a cDNA library of Norton
grape young leaves (Vitis. aestivalis var. Norton). Norton
grapevines were grown under normal greenhouse conditions.
The cDNA synthesis and library construction was performed
according to the instruction manual for pBluescript II XR
cDNA library construction kit provided by Stratagene."

ORIGIN

Query Match 35.8%; Score 321.4; DB 14; Length 782;
Best Local Similarity 73.7%; Pred. No. 5.7e-41;
Matches 409; Conservative 0; Mismatches 146; Indels 0; Gaps 0;
QY 111 AAGAAATCCCTGTTTGTGATTTGATTTGATTTGTTTGTCCAGAGCTTAGACATTTCTCGGTTT 170
Db 65 AAGAAATGCTCTCATAGACTGCATTTGATCGGCCCCCAATCTCAAGATTTCTCGAAATT 124
QY 171 CTCACCTTTAACACCGCGGTCAAGCGCGTAAGATATGCCCGCTGCTTCTCTGTGATTAATG 230
Db 125 TCCAGTTTAATTTGTGCTTAAACCGCTCAAGAAATGATGATCTTCTTCCCATTTGTCAGAG 184
QY 231 CCCAAAGCTCGCTGACTACATTCCTGATGCTAAATTTCTCAAAAGTGAAGCAATTTCTCA 290
Db 185 CTCAGAGCTCTCCAGATTATACCCAGACGCGCAGTTTACAAAGTTGAAGCGGATCTCTGA 244
QY 291 GGCCTCTGGGAGTCTCGCAAGTTTCTCGGCTTTGCTAAATAATTTGGTATTCGAGGTGTTA 350
Db 245 GGCCTCTGGGAGTCTCGGAGGTTTCTTCTCGGCTTTGCTGAAAATGGGTATTCGTGGTGTTA 304
QY 351 CTGTTTCTGATGTTGAGGTTTTTGGTGTCTCAAGTGGTTCAACTGAGAGGCGAGGCGGCT 410
Db 305 CTGTTTCTGATGTTGAGGTTTTTGGTGTCTCAAGGCGGATCGCTGAAAGACAGGCAGGCT 364
QY 411 CAGAATTTTCTGAAGACAAGTTTGTGTTGCTAAAGTTAAGATCGAGATCGTGGTTAGCAAG 470
Db 365 CTGAATTTTCTGAGGACAATTTGTTGTTGTTAAAGTTAAATGGAGATTGTGGTGAGCAAG 424
QY 471 ACCAGGTTGAGGATGTTATAGAAAAATCAATTGAGGAGGCAAGAACTGGAGAGATTGGAG 530
Db 425 ACCAGGTTGAAGCAGTAATTGACAAGATCAATTGAGGTTGGCAAGGACTGGAGAGATTGGTG 484
QY 531 ACCGCAAGATTTTCTTCTGCTGCTGTTTTCAGATGTAATAGAGTCCGCACTGGTGAGCGGG 590
Db 485 ATGCAAGATCTTTTGTGGTCCCATATCGGATGTGATAAGAGTTCGACCGGTGAACGCTG 544
QY 591 GTGATAAGGCTGAGAGGATGACAGGAGGGCGGATCTGACATGAGTACTTCTGTTGACTGC 650
Db 545 GAGAGATGGCTGAGAGGATGACTGTTGGGCGGAACCTGATATGTCCACTACAACACCGCTG 604
QY 651 TGTGACCAGCAATAT 665
|||

Db	605	CTTGAGTCCAAATAT	619	
RESULT 3				
BQ798001/c				
LOCUS				
DEFINITION				
EST 6939 Ripening Grape berries Lambda Zap II Library Vitis				
vinifera cDNA clone RT093C02 3', mRNA sequence.				
ACCESSION				
BQ798001				
VERSION				
BQ798001.1				
KEYWORDS				
EST.				
SOURCE				
Vitis vinifera				
ORGANISM				
Vitis vinifera				
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
rosids; Vitaceae; Vitis.				
REFERENCE				
1 (bases 1 to 723)				
Abbal,P., Agasse,A., Ageorges,A., Atanassova,R., Barrieu,F.,				
Couture,C., Dedaldechamp,F., Delrot,S., Glissant,D., Grimplet,J.,				
Hamdi,S., Romieu,C. and Terrier,N.				
TITLE				
Generation of Expressed Sequence Tag from Grape Berry (skin, pulp				
or seeds) at Various Developmental Stages				
JOURNAL				
Unpublished (2002)				
COMMENT				
Contact: Romieu C.				
Unite de Recherche des Produits de la Vigne				
Institut National de la Recherche Agronomique				
2, place Viala, 34 060 Montpellier Cedex 01, France				
Tel: 00-33-(0)4-99-61-28-62				
Fax: 00-33-(0)4-99-61-28-57				
Email: romieu@ensam.inra.fr				
Seq primer: T7.				
FEATURES				
Location/Qualifiers				
1..723				
/organism="Vitis vinifera"				
/mol_type="mRNA"				
/cultivar="Shiraz"				
/db_xref="taxon:29760"				
/clone="RT093C02"				
/dev_stage="ripening stage"				
/clone_lib="Ripening Grape berries Lambda Zap II Library"				
/note="Organ: Fruit; Vector: Lambda Zap II; Site_1: Eco				
RI; Site_2: XhoI; Oriented library, construction described				
in Generation of ESTs from grape Berry (skin, pulp or				
seeds) at various developmental stages by Terrier,N.				
Ageorges,A., Abbal,P., Romieu,C. in J. Plant Physiol. 158				
(12): 1575-83 2001"				
ORIGIN				
Query Match				
Best Local Similarity				
Matches				
371; Conservative				
0; Mismatches				
117; Indels				
0; Gaps				
0; Length				
723;				
QY	152	CTTAGACATTCTCGGTTTCTCACTTTAACACCGCGGTCAAGCGCGTAAGATATGCCCC	211	
Db	723	CTCAGAGATTCTCGAAATTTCAGTTAATTTGTGCCTAAACCGCTCAAAGAATGGATCT	664	
QY	212	GTGTTCTCTGTGATTAAATGCCCAAGCTCGCCTGACTACATTCTCTATGCTAAATCTAC	271	
Db	663	GTCTTCTCCATTGTTCAGAGCTCAGAGCTCTCCAGATTATACCCAGACGCGCAGTTTAC	604	
QY	272	AAAGTGAAGCAATTCTCAGGCCCTCGCGAGTCTCGCAAGTTTCTCGGTTTGTCTAAAA	331	
Db	603	AAAGTTGAAGCGATCCTGAGGCCCTGGCGAATCCAGCAGTTTCTTCGGCTTGTCTGAA	544	
QY	332	ATTGSTATTCGAGGTGTACTGTTTCTGATGTTTCGAGTTTGTGGTCTCAAGTGGTTCA	391	
Db	543	ATGGGTATTCGTGTGTACTGTTTCTGATGTCCGGGGCTTTGGTGTCTCAAGGCGGTTTG	484	
QY	392	ACTGAGAGCGAGCGGCTCAGAAATTTCTGAAGACAAGTTTGTTCGCTAAAGTTAAGATG	451	
Db	483	CCTGAAGACAGCGCAGGCTCTGAAATTTCTGAGGACAAATTTGTTGTCTAAAGTCAAAATG	424	
QY	452	GAGATCGTGGTTAGCAAAGACCAGGTTGAGGATGTTATAGAAAAATCATTGAGGAGGCA	511	

Db	423	GAGATTGTGTGAGCAAGACCAGGTTGAAGCAGTAATTGACAAGATCAATGAGTGGCA	364	
QY	512	AGAACTGGAGAGATTGGAGACGGCAAGATTTTCTTCTGCTGCCTGTTTCAGATGTAATAAGA	571	
Db	363	AGGACTGGAGAGATTTGGTGTATGGCAAGATCTTTTGGTGCCCATATCGGATGTGATAGA	304	
QY	572	GTCCGCACCTGTGAGCGGGGTGATAAGGCTGAGAGGATGACAGGAGGGCGATCTGACATG	631	
Db	303	GTTCGGACCCGTGAACGTGGAGAGCTGGCTGAGAGGATGACGGTGGCGAACTGATATG	244	
QY	632	AGTACTTC	639	
Db	243	TCCACTAC	236	
RESULT 4				
BE239550				
LOCUS				
DEFINITION				
EST403599 MHRP- Medicago truncatula cDNA clone pmHRP-28F3, mRNA				
sequence.				
ACCESSION				
BE239550				
VERSION				
BE239550.1				
KEYWORDS				
EST.				
SOURCE				
Medicago truncatula (barrel medic)				
ORGANISM				
Medicago truncatula				
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;				
Medicago.				
REFERENCE				
1 (bases 1 to 597)				
Harrison,M.J., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S.,				
Holt,I.E., Cho,J. and Fraser,C.M.				
ESTs from phosphate-starved roots of Medicago truncatula				
Unpublished (2000)				
CONTACT: Maria J. Harrison				
Plant Biology Division				
The Samuel Roberts Noble Foundation				
2510 Sam Noble Parkway, Ardmore, OK 73401, USA				
Tel: 580-223-5810				
Fax: 580-221-7380				
Email: mjharrison@noble.org				
The Samuel Noble Roberts Foundation: N265417e				
TIGR sequence name:MTHAC26TK				
More information is available at..				
http://chrysie.tamu.edu/medicago				
Seq primer: SKmod (CTA gAA CTA gtg gAT CC).				
Location/Qualifiers				
1..597				
/organism="Medicago truncatula"				
/mol_type="mRNA"				
/cultivar="A17"				
/db_xref="taxon:3880"				
/clone="pmHRP-28F3"				
/tissue_type="roots"				
/dev_stage="phosphate-starved"				
/lab_host="XL0LR"				
/clone_lib="MHRP"				
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:				
XhoI; At the trifoliolate stage, M. truncatula plants were				
transplanted to phosphate-free sand and grown for a				
further 30 days. During this period, they were fertilized				
twice weekly with 1/2 Hoaglands solutions containing 20uM				
potassium phosphate. cDNA was prepared from polyA+				
enriched RNA. The cDNA was directionally ligated into the				
Unizap XR vector from Stratagene and packaged using				
Gigapack III Gold packaging extracts. Plasmids containing				
cDNA inserts were excised from the recombinant lambda-Zap				
phage using Ex-assist helper phage and propagated in				
XL0LR cells."				
ORIGIN				
Query Match				
33.3%; Score 298.6; DB 10; Length 597;				

Best Local Similarity 71.8%; Pred. No. 2.5e-37; Mismatches 154; Indels 6; Gaps 1;
Matches 407; Conservative 0;

QY 62 ACTGCGAACTGGGCTTGGCTCACTCTCTTCAATCTAATAACATCAAGAAAGAAATTCCT 121
Db 28 ATTGCGAAACCGAACGCTTCAACGGTTTGAATTTTCATATCAATGAAGAACTCAATTTCT 87

QY 122 GTTTTGTATTTTCACTTTTGTTCCTCAGAGCTTAGACATTTCTCGGTTTCTCACTTTAAC 181
Db 88 TTTTCAAGCTTCAGCGTCATTCGCAAGCGTTTCGGAGATTC-----TTCTCATCGCAAT 141

QY 182 ACCGCGTCAAGCGGTAAGATATGCCCCCGTCTGTTCTCTGTGATTAAATGCCCAAAGCTCG 241
Db 142 GTGTCCTAAAGCAACGGAATGTCATGCTTCTCCCAAAATCAGAGCTCAAAACCTT 201

QY 242 CTTGACTACATTCCTGATGCTAAATTTCTACAAAGTGAAGCAATTTCTCAGGCCCTGGGA 301
Db 202 CTTGACTATGTTCTGATCCAGTTTACAAAGTTGAAGCCATTTCTCAGGCCATGGAGA 261

QY 302 GTCTCGCAAGTTTCTCGGCTTTCTTAAAAATTTGGTATTCGAGTGTACTGTTCTGAT 361
Db 262 ATCCCTCAGGTTTCTTCGGGTTTGTGAAAATGGGAATTCGTGTTCTCACTGATCTGAT 321

QY 362 GTTCGAGGTTTGTGCTCAAGGTTTCAACTGAGAGCGAGGCGGCTCAGAAATTTTCT 421
Db 322 GTCAAGGGTTTGTGCTCAGGTTGCTCAAAAGAGAGGCGAGGCTCCGAATTTTCT 381

QY 422 GAAGACAAGTTTGTGCTAAAGTTAAGATCGAGATCGTGTAGCAAGAACAGGTTGAG 481
Db 382 GAAGACAATTTTGTGCCAAAGTTAAATGGAATAGTGGTGAGAAAAGACCAGGTTGAG 441

QY 482 GATGTTATAGAAAAATCATTTGAGGAGGCGAGAACTGGAGAGATTGGAGACGGCAAGATT 541
Db 442 GCAGTGATAAACAAAAATTTATGGAACGCGGAAGAACTGGGAGATTGGTATGGCAAAAT 501

QY 542 TTCTTGCTGCTGTTTTCAGATGTAATAAGACTCCGCACTGGTGAGCGGGGTGATAAGGCT 601
Db 502 TTCTTGATCCCTGTATCTGATGTAATAAGAAATCCGACAGGTTGAGCGGGAGCAGGCT 561

QY 602 GAGAGGATGACAGAGGCGGATCTGAC 628
Db 562 GAGAGGATGGCTGGGGGACTAACTGAC 588

RESULT 5
BI311072
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BI311072
EST5312822 GSD Medicago truncatula cDNA clone pGESD9J4 5' end,
mRNA sequence.
BI311072
BI311072.1 GI:14985399
EST.
Medicago truncatula (barrel medic)
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 714)
Grusak, M.A., Samac, D.A., Town, C.D., Van Aken, S., Utterback, T.,
Cho, J. and Fraser, C.M.
ESTs from developing reproductive tissues of Medicago truncatula
Unpublished (2001)
Contact: Michael A. Grusak
USDA/ARS Children's Nutrition Research Center
Baylor College of Medicine
1100 Bates Street, Houston, TX 77030-2600, USA
Tel: 713-798-7044
Fax: 713-798-7078
Email: mgrusak@bcm.tmc.edu
B397927e

TIGR sequence name: MTPAP50TK
More information is available at: www.medicago.org

Seq primer: SKmod (CTA GAA CTA gtg gAT CC).
Location/Qualifiers
1. 714
/organism="Medicago truncatula"
/mol_type="mRNA"
/culturivar="A17"
/db_xref="taxon:3880"
/clone="pGESD9J4"
/tissue_type="Immature seeds"
/dev_stage="Immature seeds, 11 to 19 days after
pollination"
/clone_lib="GESD"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Immature seeds, collected from pods ranging in age
from 11 to 19 days after pollination, were harvested from
greenhouse-grown plants. Seeds were removed and
separated from pod walls and were immediately frozen in
liquid nitrogen. Seeds throughout the age range were
pooled for mRNA extraction. cDNA was prepared from polyA+
enriched RNA. The cDNA was directionally ligated into
the Unizap XR vector from Stratagene and packaged using
Gigapack III Gold packaging extracts. Plasmids containing
cDNA inserts were excised from the recombinant lambda-Zap
phage using Ex-assist helper phage and propagated in
XL01R cells."

FEATURES
source

ORIGIN

Query Match 33.3%; Score 298.5; DB 12; Length 714;
Best Local Similarity 71.8%; Pred. No. 2.2e-37;
Matches 407; Conservative 0; Mismatches 154; Indels 6; Gaps 1;

QY 62 ACTGCGAACTGGGCTTGTCTCACTCTCTTCAATTTCTAATAACATCAAGAAAGAAATTCCT 121
Db 27 ATTGCGAAACCGAACGTTTCAACCGTTTGAATTTTCATATCAATGAAGAACTCAATTTCT 86

QY 122 GTTTTGTATTTTGTGTTTGTCCAGAGCTTAGACATTTCTCGGTTTCTCACTTTAAC 181
Db 87 TTTTCAAGCTTCAGCGTCATTCGCAAGCGTTTCGGAGATTC-----TTCTCATCGCAAT 140

QY 182 ACCGCGTCAAGCGGTAAGATATGCCCGCTCGTTCCTGTGATTAAATGCCCAAAGCTCG 241
Db 141 GTGGTCTTAAAGCAACGGAATGTCATCGATTTCTCCAAATCAGAGCTCAAAACCTT 200

QY 242 CCTGACTACATTTCTGATGCTAAATTTCTACAAAGTGAAGCAATTTCTCAGGCCCTGGGA 301
Db 201 CCTGACTATGTTCTGATCCAGTTTTCACAAAGTTGAAGCCATTTCTCAGGCCATGGAGA 260

QY 302 GTCTCGCAAGTTTCTCGGCTTGTCTTAAAAATTTGGTATTCGAGGTTGTTACTGTTCTGAT 361
Db 261 ATCCCTCAGGTTTCTTCGGGTTTGTGAAAATGGGAATTCGTGTTCTCACTGATCTGAT 320

QY 362 GTTCGAGGTTTGTGCTCAAGGTTGTTCTCACTGAGAGCGAGGCGGCTCAGAAATTTTCT 421
Db 321 GTCAAGGGGTTTGTGCTCAGGTTGGTCTCAAAAGAGAGCGAGGCTCCGAATTTTCT 380

QY 422 GAAGACAAGTTTGTGCTTAAAGTTAAGATGGAATCGTGTAGCAAGAACCGATTGAG 481
Db 381 GAAGACAATTTTGTGCCAAAGTTTAAATGGAATAGTGGTGAGAAAAGACCAGGTTGAG 440

QY 482 GATGTTATAGAAAAATCATTTGAGGAGGCGAAGAACTCGAGAGATTGGAGACGGCAAGATT 541
Db 441 GCAGTGATAAACAAAAATTTATGAGACGGCAAGAACTCGGGAGATTGGTATGGCAAAAT 500

QY 542 TTCTTGCTGCTGTTTTCAGATGTAATAAGAGTCCGCACTGGTGAGCGGGGTGATAAGGCT 601
Db 501 TTCTTGATCCCTGTATCTGATGTAATAAGAAATCCGACAGGTTGAGCGGTGGGGAGCAGGCT 560

QY 602 GAGAGGATGACAGAGGCGGATCTGAC 628
Db 561 GAGAGGATGGCTGGGGGACTAACTGAC 587

RESULT 6

CA922829/c
LOCUS
DEFINITION CA922829 812 bp mRNA linear EST 09-MAY-2003
EST640547 MTUS Medicago truncatula cDNA clone MTUS-58H7, mRNA
sequence.
ACCESSION
VERSION CA922829.1 GI:27409759
KEYWORDS
SOURCE
ORGANISM
Medicago truncatula (barrel medic)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 812)
VandenBosch,K., Endre,G., Silverstein,K, Town,C.D., Van Aken,S.,
Utterback,T., Cheung,F. and Fraser,C.M.
The Medicago truncatula 6K unigene set: cDNA clones selected and
re-arrayed from various libraries
Unpublished (2002)
Contact: VandenBosch K
Department of Plant Biology
University of Minnesota
220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
Tel: 612 624 2755
Fax: 612 625 1738
Email: kvandenb@chbbs.umn.edu
Alias Clone pMHRP-28F3
TIGR sequence name: MTUCF91TV
More information is available at: www.medicago.org
Seq primer: (gTA AtA CgA CtC Act AtA ggg C).
Location/Qualifiers
1 .812
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="A17"
/db_xref="taxon:3880"
/clone="MTUS-58H7"
/tissue_type="mixed tissues"
/dev_stage="various stages"
/lab_host="XLOLR"
/clone_lib="MTUS"
/note="vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in XLOLR cells."
ORIGIN
Query Match 33.1%; Score 297; DB 14; Length 812;
Best Local Similarity 71.6%; Pred. No. 3.6e-37;
Matches 406; Conservative 0; Mismatches 155; Indels 6; Gaps 1;
62 ACTGCGAAACTGGGCTTGCTCACTCCTCTTCATTTCTAATAACATCAAGAAAGAAATTCCT 121
|||||
768 ATTGCGAAACCGAACGGTGTCAACGGTTGAAATTTTCATATCAATGAAGAACTCAATTCCT 709
|||||
122 GTTTTGTATTTCAGTTTGTGTTGTCAGAGCTTAGACATTTCTGGTTTCTCATTAAAC 181
|||||
708 TTTTCAAGCTTCAGCGTCAATTCGCAAGCGTTTCGGAGATTC-----TTCTCATCGCAAT 655
|||||
182 ACCGCGGTCAAGCGGTAAAGATATGCCCCCGTTCGTTCTGTGATTAATGCCCAAGCTCG 241
|||||
654 GTGGTCCTAAAAGCAACGGAATGTCATCGATTCTTCCAAAATCAGAGCTCAAAACCTT 595
|||||
242 CCTGACTACATTCCTGATGCTAAATTTACAAAGTGAAGCAATTTCTCAGGCCCTGGCGA 301
|||||
594 CCTGACTATGTTCTCTGATCCAGTTTACAAAGTTGAAGCCATTTCTCAGGCCATGGAGA 535
|||||
302 GTCTCGCAAGTTTCCTCGGCTTTGCTAAATAATGGTATTCGAGGTGTTACTGTTCTGAT 361
|||||
534 ATCCCTCAGGTTTCTTCGGGTTTGTGTAATGGGAATTCGGTGTCACTGTATCTGAT 475
|||||

QY 362 GTTCGAGGTTTGTGCTCAAGGTGGTTCAACTGAGAGCGAGGGCGGCTCAGAAATTTTCT 421
|||
Db 474 GTCAAGGGGTTTGTGCTCAGGGTGGCTCAAAGAGAGAGCGAGGGAGGCTCCGAATTTTCT 415
|||
QY 422 GAAGACAAAGTTTGTGCTAAAGTTAAGATGGAGATCGTGTTAGTACAAAGACCGAGTTGAG 481
|||||
Db 414 GAAGACAAATTTTGTGCCAAAGTTAAATAGTAAATAGTGTGAGAAAGACCGAGTTGAG 355
|||||
QY 482 GATGTTATAGAAAAATCAATTGAGGAGGCAAGAACTGGAGAGATTGGAGACGGCAAGATT 541
|||||
Db 354 GCAGTGATAAACAAATTTATGGAGACGGCAAGAACTGGGAGATTGGTGATGGCAAAATT 295
|||||
QY 542 TTCTTGCTGCTGCTTTCAGATGTAATAAGAGTCCGCACTGTGTGAGCGGGTGATAAGGCT 601
|||||
Db 294 TTCTTGATCCCTGTATCTGTATGATTAATAAGATCCGCACAGGTGAGCGTGGGAGCAGGCT 235
|||||
QY 602 GAGAGGATGACAGGAGGGCGGATCTGAC 628
|||||
Db 234 GAGAGGATGGCTGGGGGACTAACTGAC 208
|||||
RESULT 7
BF645522 612 bp mRNA linear EST 20-DEC-2000
LOCUS
DEFINITION NF036F08EC1F1074 Elicited cell culture Medicago truncatula cDNA
clone NF036F08EC 5', mRNA sequence.
ACCESSION BF645522
VERSION BF645522.1 GI:11910651
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 612)
Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation -
Center for Medicago Genomics Research
Unpublished (2000)
Contact: Dixon RA
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7302
Fax: 580 221 7380
Email: radixon@noble.org
Insert Length: 612 Std Error: 0.00
Plate: 036 row: F column: 08
Seq primer: TCACACAGGAAACAGCTATGAC.
Location/Qualifiers
1 .612
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF036F08EC"
/tissue_type="Cell cultures derived from root tissues"
/dev_stage="Cell suspensions were subcultured every 14
days. Cells were induced six days after subculture"
/clone_lib="Elicited cell culture"
/note="vector: Lambda Zap; Cells were induced with yeast
cell wall extracts equivalent to 50ug/ml glucose in the
final concentration. Samples were taken at 0.5, 1, 12 and
24 hours after induction. Equal amounts of RNA from each
time point were pooled and used for mRNA isolation."
ORIGIN
Query Match 32.7%; Score 293; DB 10; Length 612;
Best Local Similarity 70.9%; Pred. No. 1.8e-36;
Matches 402; Conservative 0; Mismatches 159; Indels 6; Gaps 1;

QY	62	ACTGCGAAACTGGGCTTGGCTCACTCCTCTTCAATCTATAATACATCAAGAAAGAAATTCCT	121
Db	44	ATTGCGAAACCGAACGTGTTCAACGGTTTGAATTTTCATATCAATGAAACTCAATTCCT	103
QY	122	GTTTTGAATTCAGTTTGTGTTGTCAGAGCTTAGACATTCCTCGGTTTCTCACTTTAAC	181
Db	104	TTTTCAAGCTTCAGCGTCAATCGCAAGCGTTTCGGAGATTC-----TTCTCATCGCAAT	157
QY	182	ACCGCGTCAAGCGGTAAGATATGCCCGCTCGTTCCTGTGATTAATGCCCAAGCTCG	241
Db	158	GTGGTCTTAAAGCAACGGAATGATCGATTTCTTCCCAAAATCAGAGCTCAGAACCTT	217
QY	242	CCTGACTACATTCCTGATGCTAAATTTCTACAAAGTGAAGCAATTTCTCAGGCCCTGGGA	301
Db	218	CCTGACTATGTTCTCTGAATCCAAAGTTTACAAAGTTGAAGCCATTTCTCAGGCCATGGGA	277
QY	302	GTCTCGAAGTTTCTCGGCTTTGCTTAAATAATGCTATTCGAGGTGTTACTGTTTCTGAT	361
Db	278	ATCCCTCAGGTTTCTTCGGGTTTGTGAAATGGAATTCGTGGTGTCACTGTATCTGAT	337
QY	362	GTTCGAGTTTGGTCTCAAGTGGTTCAACTGAGAGCGGCGGCTCAGAAATTTTCT	421
Db	338	GTCAAGGGGTTTGGTCTCAGGGTGGCTCAAAAGATAGGCAGGGAGGCTCCGAATTTTCT	397
QY	422	GAAGACAAGTTTGTGCTAAAGTTAAGATGAGATCGTGGTTAGCAAAAGACAGGTTGAG	481
Db	398	GAAGACAATTTTGTCCCAAGTTAAATGGAATAGTGGTGAGAAAGACAGGTTGAG	457
QY	482	GATGTTATAGAAAAATCATTTGAGGAGGCGCAAGAACTGGAGAGATTGGAGACGCGCAAGATT	541
Db	458	GCAGTGATAAACAAATATATGGAGACNGCAAGAACTGGGGAGATTGGTGATGGCAAAATT	517
QY	542	TTCTTGCTGCTGTTTCAGATGTAATAAGATCCGCACTGGTGAGCGGGGTGATAAGGCT	601
Db	518	TTCTTGATCCCTGTATCTGATGTAATAAGAAATCCNACAGGTGAGCGTGGGGAGCANGCT	577
QY	602	GAGAGGATGACAGGAGCGGCGATCTGAC	628
Db	578	GANAGGATGGCTGGGGGACTAACTGAC	604
RESULT 8			
CF920467			
LOCUS			
DEFINITION			
gmhRw3-14_G03_1_019 Soybean root hair subtracted cDNA library			
gmhRw3 Glycine max cDNA, mRNA sequence.			
CF920467			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Glycine max (soybean)			
Glycine max			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;			
Glycine.			
1 (bases 1 to 759)			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Unpublished (2003)			
Contact: Gary Stacey			
University of Missouri			
108 Waters Hall, Columbia, MO 65211, USA			
Tel: 573-884-4752			
Fax: 573-882-0588			
Email: staceyg@missouri.edu			
Single pass sequence			
Seq primer: T7.			
Location/Qualifiers			
1..759			
/organism="Glycine max"			
/mol type="mRNA"			
/cultivar="Williams 82"			
FEATURES			
Source			
1..759			
/organism="Glycine max"			
/mol type="mRNA"			
/cultivar="Williams 82"			
RESULT 9			
C81819			
LOCUS			
DEFINITION			
C81819 Citrus unshiu juice sac and pulp segment maturation stage			
Citrus unshiu cDNA clone pcMRM01.17-061, mRNA sequence.			
C81819			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Citrus unshiu			
Citrus unshiu			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
rosids; eurosids II; Sapindales; Rutaceae; Citrus.			
1 (bases 1 to 866)			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Mitsuo Omura			
Department of Citrus Research			
National Institute of Fruit Tree Science, Okitsu			
Okitsu 485-6, Shimizu, Shizuoka 424-0292, Japan			
Tel: 81-543-69-7108			
Fax: +81-543-69-2115			
FEATURES			
Source			
1..866			
/organism="Citrus unshiu"			
/mol type="mRNA"			
/cultivar="Unshiu"			
/segment="juice sac and pulp segment maturation stage"			
/clone="pcMRM01.17-061"			
/molecule="mRNA"			
/sequence="mRNA sequence"			
/date="1999-01-17"			
/accession="C81819"			
/version="1"			
/keywords="Citrus unshiu"			
/source="Citrus unshiu"			
/organism="Citrus unshiu"			
/mol type="mRNA"			
/cultivar="Unshiu"			
/segment="juice sac and pulp segment maturation stage"			
/clone="pcMRM01.17-061"			
/molecule="mRNA"			
/sequence="mRNA sequence"			
/date="1999-01-17"			
/accession="C81819"			
/version="1"			
/keywords="Citrus unshiu"			
/source="Citrus unshiu"			
/organism="Citrus unshiu"			
/mol type="mRNA"			
/cultivar="Unshiu"			
/segment="juice sac and pulp segment maturation stage"			
/clone="pcMRM01.17-061"			
/molecule="mRNA"			
/sequence="mRNA sequence"			
/date="1999-01-17"			
/accession="C81819"			
/version="1"			
/keywords="Citrus unshiu"			
/source="Citrus unshiu"			
/organism="Citrus unshiu"			
/mol type="mRNA"			
/cultivar="Unshiu"			
/segment="juice sac and pulp segment maturation stage"			
/clone="pcMRM01.17-061"			
/molecule="mRNA"			
/sequence="mRNA sequence"			
/date="1999-01-17"			
/accession="C81819"			
/version="1"			
/keywords="Citrus unshiu"			
/source="Citrus unshiu"			
/organism="Citrus unshiu"			
/mol type="mRNA"			
/cultivar="Unshiu"			
/segment="juice sac and pulp segment maturation stage"			
/clone="pcMRM01.17-061"			
/molecule="mRNA"			
/sequence="mRNA sequence"			
/date="1999-01-17"			
/accession="C81819"			
/version="1"			
/keywords="Citrus unshiu"			
/source="Citrus unshiu"			
/organism="Citrus unshiu"			
/mol type="mRNA"			
/cultivar="Unshiu"			
/segment="juice sac and pulp segment maturation stage"			
/clone="pcMRM01.17-061"			
/molecule="mRNA"			
/sequence="mRNA sequence"			
/date="1999-01-17"			
/accession="C81819"			
/version="1"			
/keywords="Citrus unshiu"			
/source="Citrus unshiu"			
/organism="Citrus unshiu"			
/mol type="mRNA"			
/cultivar="Unshiu"			
/segment="juice sac and pulp segment maturation stage"			
/clone="pcMRM01.17-061"			
/molecule="mRNA"			
/sequence="mRNA sequence"			
/date="1999-01-17"			
/accession="C81819"			
/version="1"			
/keywords="Citrus unshiu"			
/source="Citrus unshiu"			
/organism="Citrus unshiu"			
/mol type="mRNA"			
/cultivar="Unshiu"			
/segment="juice sac and pulp segment maturation stage"			
/clone="pcMRM01.17-061"			
/molecule="mRNA"			
/sequence="mRNA sequence"			
/date="1999-01-17"			
/accession="C81819"			
/version="1"			
/keywords="Citrus unshiu"			
/source="Citrus unshiu"			
/organism="Citrus unshiu"			
/mol type="mRNA"			
/cultivar="Unshiu"			
/segment="juice sac and pulp segment maturation stage"			
/clone="pcMRM01.17-061"			
/molecule="mRNA"			
/sequence="mRNA sequence"			
/date="1999-01-17"			
/accession="C81819"			
/version="1"			
/keywords="Citrus unshiu"			
/source="Citrus unshiu"			
/organism="Citrus unshiu"			
/mol type="mRNA"			
/cultivar="Unshiu"			
/segment="juice sac and pulp segment maturation stage"			
/clone="pcMRM01.17-061"			
/molecule="mRNA"			
/sequence="mRNA sequence"			
/date="1999-01-17"			
/accession="C81819"			
/version="1"			
/keywords="Citrus unshiu"			
/source="Citrus unshiu"			
/organism="Citrus unshiu"			
/mol type="mRNA"			
/cultivar="Unshiu"			
/segment="juice sac and pulp segment maturation stage"			
/clone="pcMRM01.17-061"			
/molecule="mRNA"			

FEATURES source
Email: xx145039@affrc.go.jp.
Location/Qualifiers
1. 866
/organism="Citrus unshiu"
/mol_type="mRNA"
/cultivar="Miyagawa-Wase"
/db_xref="taxon:55188"
/clone="pcMFRM01.17-061"
/tissue_type="juice sac and pulp segment"
/dev_stage="maturation stage"
/clone_lib="Citrus unshiu juice sac and pulp segment
maturation stage"
ORIGIN
Query Match 31.5%; Score 283; DB 13; Length 866;
Best Local Similarity 71.8%; Pred. No. 5.3e-35;
Matches 484; Conservative 0; Mismatches 143; Indels 47; Gaps 7;
QY 8 CGGCCGCTTAGAAGTGGATCCCGGGCTGCAGGAATTCGGCAGGAG----- 58
Db 20 CGNCGCTTAGAAGTGGATCCCGGGCTGCAGGAATTCGGCAGGATCGTTAA 79
QY 59 -----GCTACTGGGAACCTGGGCTTGCTCACTCCTCTTCATTCTA 98
Db 80 CTGGAGGACCGTAATGGCTGCAATTGCAACCCAGCCTCGTTACTCCGCTCAACTTCT 139
QY 99 ATAACATCAAGAAAGAAATCCCTGTTTGTGATTTTCACTTGTGTTTGTCCAGAGCTTAGAC 158
Db 140 CTTCTCATCTCAAGAACTTCCTCTTGTAATTTCCGCTCCGTCGGCCAAAGTTTGAGAA 199
QY 159 ATTCTCGGTTTCTCACTTTAAACACCGCGGTCAAGCGGTAAGATATGCCCGCTCGTTTC 218
Db 200 TTTCTCGACTTTCTAG-----GCTCAAAATGCACAAGAAATGTATCAATCTTC 247
QY 219 CTGTGATTAATGCCAAAGCTCGCTGACTACATTCCTGATGCTAAATTTTACAAAGTGG 278
Db 248 ATGTTATCAGAGCCCGAGAGCTCACCTGATTAATTCCTGATTTAAGTTTACAAAGTGG 307
QY 279 AAGCAATTCTCAGGCCCTCGCGAGTCTCGCAAGTTTCTCGGCTTGCTAAAAATTCGTA 338
Db 308 AAGCCATTCTCAGGCCCTCGGAGAGTCCAGCAGGTTTCTCGGCTTGCTGAATATGGA 367
QY 339 TTCGAGGTGTTACTGTTTCTGATGTTTCTGA-GGTTTGGTGCTCAAGGTGGTTCAAC-TGA 396
Db 368 TCCGTGGTGTACGGTTTCTGATGTTTCTGATGTTTGGTGCTCAAGGTGGTCAACAGGA 427
QY 397 GAGGCAGGGCGGCT--CAGAAATTTCTGA-AGACAAGTTTGTGCTAAAGTTAAGATGGA 453
Db 428 AAGGCATGTTGGTCTCCCTGAATTTTCTGAGGGGAAATTTGTCNAAAGTGAAGATGGA 487
QY 454 GATCGTGGTTAGCAAGACCAAGCTGAGGATGTTATAGAAAAAATCATTGAGGAGGCAAG 513
Db 488 GATTGTGG-TAACAAAGACCAAGGTGAAGGAGTAATTGATAAAATTTATGGAGGAGGCANG 546
QY 514 AACTGGAGAGATTGGAGACGCGCAAGATTTTCTTGCTGCCCTGTTTCAGATGTAATAAGAGT 573
Db 547 GACTGGAGAAATTGGTGACGCGCAAGATTTTGTGGTACCTGTCTCAGATGTCTAAGAGT 606
QY 574 CCGCACTGGTGAGCGGGCTGATAAGGCTGAGAGGATGACAGGAGGCGGATCTGACATGAG 633
Db 607 TCGAACTGGTGAACGAGGTGAAGAAAGCTGAAAGGATGGCAGGAGGATGGTCGGACATATC 666
QY 634 TACTTCTGCTTGAC 647
Db 667 TTCAGCTGTGTGAC 680
RESULT 10
BE823431/c
LOCUS
DEFINITION GM700019B23E12 Gm-r1070 Glycine max cdna clone Gm-r1070-7536 3',
mRNA sequence.
ACCESSION BE823431

VERSION BE823431.1 GI:10255665
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1 (bases 1 to 677)
AUTHORS Vodkin,L., Keim,P., Shoemaker,R., Retzel,E., Khanna,A., Coryell,V.,
Erpelting,J., Raph,C., Shoop,E., Pardinas,J., Liu,L. and Lewin,H.
TITLE A Functional Genomics Program for Soybean (NSF 9872565)
JOURNAL Unpublished (1999)
COMMENT Other ESTs: AW471508 corresponding to Gm-cl029-983 (5')
Contact: Vodkin, L.O., PI, A Functional Genomics Program for
Soybean (NSF 9872565)
Lewin, H. A., Director, Keck Center for Comparative and Functional
Genomics
University of Illinois
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
Tel: (217) 244-6147
Fax: (217) 333-4582
Email: l-vodkin@uiuc.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134. For further information
call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
427-3324 or contact:clones@genomesystems.com or info@genome
systems.com web site:www.genomesystems.com
Seq primer: 5'-TTTTTTTCTTTTCTTTT(A/C/G)-3'.
FEATURES Location/Qualifiers
1. 677
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="Gm-r1070-7536"
/clone_lib="Gm-r1070"
/note="The library Gm-r1070 is a sequence-driven, reracked
set of 9,216 clones selected from cDNA libraries from
various tissues and stages of development of soybean that
represent 2,639 sequences from immature cotyledons, 1,770
from immature seed coats, 3,938 from flowers, and 869
from young pods. The 5' ESTs of the source clones from
the different libraries was used to select singletons, or
a representative of each contig, which were reracked to
form library Gm-r1070. The cDNA clones of the reracked
Gm-r1070 library were then sequenced at the 3' end. The
contig analysis to select unique genes was performed by
the laboratory of Ernest Retzel, Center for Computational
Genomics and Bioinformatics, University of Minnesota.
http://www.cbc.umn.edu/ResearchProjects/Soybean/index.html
. Reracking was performed by Genome Systems, St. Louis,
http://www.genomesystems.com, and 3' sequencing by the
Keck Center for Comparative and Functional Genomics,
University of Illinois.
http://www.life.uiuc.edu/biotech/keck.html. Note: The
corresponding 5' EST from each clone in the Gm-r1070
library is listed in the 'OTHER EST' field. The detailed
information on the source library for each clone can also
be obtained by referring to the Genome Systems clone ID of
the original cDNA library that is also listed under
'OTHER EST'.
ORIGIN
Query Match 31.0%; Score 277.8; DB 10; Length 677;
Best Local Similarity 72.2%; Pred. No. 4e-34;
Matches 345; Conservative 0; Mismatches 133; Indels 0; Gaps 0;
QY 194 CGCGTAGATATGCCCCGTCGTTCTCTGTGATTAATGCCAAAGCTCGCTGACTACATT 253
Db 675 CNCAGAGTAAATNNNNNNNNNTTCCCCAAATCAGAGCCCCAGNNNNNTCCAGATTATGTT 616
QY 254 CCTGATGCTAAATTTCTACAAAGTGGAGCAATTTCTCAGGCCCTGGCGAGTCTCGAAGTT 313

Db 615 CCAAATCCGAGTTTACAAAGTAGAGCGGATTCTCAGNNCATGGCGAGTTCNNNNGGTT 556

Qy 314 TCCTCGGCTTGCTAAATAATGGTATTCGAGGTGTTACTGTTTCTGATGTTTCGAGGTTT 373

Db 555 TCTCGGCTTGTTGAAATGGGAATTCGGTGTCACTGTATCTGATGTTCAGGGGCTTT 496

Qy 374 GGTGCTCAAGGTGGTTCAACTGAGAGGCGGGGGCTCAGAAATTTCTGAAGACAAGTTT 433

Db 495 GGTGCTCAGGTGGTTCAAAAGAGAGGCGAGGCTCCGAATTTTCAGAGACAATTTT 436

Qy 434 GTTGCTAAAGTTAAGATGGAGATCGTGGTTAGCAAGACCAGGTTGAGGATGTTATAGAA 493

Db 435 GTTGCCAAAGTTAAATGGAAGTAGTGGTGAGAAAGGACCAGGTTGAGGCAAGTTATTGAC 376

Qy 494 AAAATCATTTGAGGAGCGCAAGAACTGGAGAGATTGGAGACGGCAAGATTTCCTGCTGCCT 553

Db 375 AAAATTTATTGAGGAGCGCAAGAACTGGGAGATTGGTGATGGCAAAATTTTCTTGATCCCC 316

Qy 554 GTTTCAGATGTAATAAGAGTCCGCACTGCTGAGCGGGGTGATAAGGCTGAGAGGATGACA 613

Db 315 ATCTCAGATGTTATAAGAATTCGGACAGGTGACCTGGGAGCGAGGCAAGCTAGGATGACT 256

Qy 614 GGAGGGCGATCTGACATGAGTACTTCTGCTTGTGCTGCTGACCGACCAATATAGCAAT 671

Db 255 GGGGGCCGAAGTGACATGTTATCTGCTGTATGACTGGTTGGAATGAAAGATAAGCAAT 198

RESULT 11

BM113473

LOCUS

DEFINITION

EST561009 potato roots Solanum tuberosum cDNA clone cPRO18M20 5'

end, mRNA sequence.

BM113473

BM113473.1 GI:17076521

EST.

Solanum tuberosum (potato)

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE

AUTHORS

van der Hoeven,R., Sun,H., Karamycheva,S.A., Tsai,J., Van Aken,S., Utterback,T., Chieningo,A., Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.

Generation of ESTs from potato roots

Unpublished (2001)

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:

http://genome.arizona.edu/orders/

Seq primer: T3.

FEATURES

source

1. .799

/organism="Solanum tuberosum"

/mol_type="mRNA"

/cultivar="Kennebec"

/db_xref="taxon:4113"

/clone="cPRO18M20"

/tissue type="roots"

/dev_stage="in vitro grown stem cuttings"

/lab_host="SOLR"

/clone_lib="potato roots"

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Cornell University, Tanksley lab; sequencing; The Institute for Genomic Research. Roots were isolated from in vitro grown stem cuttings on CM medium. Roots were isolated two weeks after placing the stem cuttings from in vitro grown plants on medium."

ORIGIN

Query Match 30.8%; Score 276; DB 12; Length 799;

Best Local Similarity 70.4%; Pred. No. 6.9e-34;

Matches 369; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

Qy 153 TTAGACATTCGCGTTTCTCACTTTAAACACCGCGGTCAAGCGTAAAGATATGCCCGG 212

Db 105 TTGTTCAACCCCAAGTTCTTCCCTTCGCAGCTCACCGTCAAAACGTTGTCAAAATGCTCCTT 164

Qy 213 TCGTTCCCTGTGATTAATGCCAAAGCTCGCTGACTACATTCCTGATGCTAAATTTACA 272

Db 165 CTATCCCGATTATCAGAGCTCAAAACCTCTCCAGATTTCTGCCGATGCAAAATTTACA 224

Qy 273 AAGTGAAGCAATTTCTCAGGCCCTGGCGAGTCTCGCAAGTTTCTCGGCTTTGCTAAAAA 332

Db 225 AAGTCGAAGCAATTTCTAGACCTTGGAGAATTTCAACAGGTTTCTTCGGCACTACTGAAA 284

Qy 333 TTGGTATTTCGAGGTGTTACTGTTTCTGATGTTTCGAGGTTTGGTCTCAAGGTGTTCAA 392

Db 285 TGGGCATTCGTGGTGTCACTGTTTTCGATGTTTCGAGGTTTGGCGCCCAAGGAGGTTGA 344

Qy 393 CTGAGAGCGAGGGCGGCTCAGAAATTTTCTGAAGACAAGTTTGTGCTAAAGTTAAGATGG 452

Db 345 CTGAGAGGCAAGCTGGCTCTGAATTTCTGAAGACACATTTGTTGCAAAAGTTAAATGG 404

Qy 453 AGATCGTGTGTTAGCAAAAGACCAAGGTTTGAGGAGTTTATAGAAAATCATTGAGGAGGCAA 512

Db 405 AAATTGTTGTCAGCAAAAGACCAAGGTTTGAAGAGTCAATTCGCAAGATAATTTGAAGAGGCAA 464

Qy 513 GAACTGGAGAGATTGGAGACGGCAAGATTCTTCTGCTGCTGTTTTCAGATGTAATAAGAG 572

Db 465 GAACTGGTGAATAGCGGATGGAAGATATTTCTTGACTCCCATCTCTGATGTAATAAGAG 524

Qy 573 TCCGCACTGCTGAGCGGGGTGATAAGGCTGAGAGGATGACAGGAGGGGATCTGACATGA 632

Db 525 TTCGCACTGGAGAACCGGGAGAAAGGACAGAGGATGATGGAGGGCATGCTGATATGT 584

Qy 633 GTACTTCTGCTGACTGCTGTGACCAAGCAATATATAGCATTCAGGA 676

Db 585 CCTCTGCATTATCAACTTCTTGAGCGGCAATGCCCTATCTTTGA 628

RESULT 12

BI932336

LOCUS

DEFINITION

EST552225 tomato flower, 8 mm to preanthesis buds Lycopersicon esculentum cDNA clone cTOC22H10 5' end, mRNA sequence.

ACCESSION

BI932336

BI932336.1 GI:16246808

VERSION

KEYWORDS

SOURCE

ORGANISM

Lycopersicon esculentum (tomato)

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE

AUTHORS

van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J., Utterback,T., Van Aken,S., Ronning,C.M., Niernan,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.

Generation of ESTs from tomato flower tissue, buds 8 mm - preanthesis

Unpublished (2001)

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

This clone is available through the Clemson University Genomics Institute

TITLE

JOURNAL

COMMENT

FEATURES

source

1. .711

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLOC22H10"
/tissue_type="flower"
/dev_stage="buds 8mm to preanthesis"
/clone_lib="tomato flower, 8 mm to preanthesis buds"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research; Flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

ORIGIN

Query Match 30.7%; Score 275.8; DB 12; Length 711;
Best Local Similarity 73.5%; Pred. No. 8e-34;
Matches 352; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 184 CGCGGTCAAGCGGTAAAGATATGCCCCCGTCGTTCTCTGTGATTAATGCCAAAGCTCGCC 243
Db 175 CACCTTCAAACGCTGTCAAAATGCTCCTTCTTCCCGATTATCAGAGCCCCAAACTCTCC 234

QY 244 TGAATACATTCCTGATGCTAAATTTCTACAAAGTGGAAGCAATTTCTCAGGCCCTCGCGAGT 303
Db 235 AGATTTCGTCCCTGATGCGAAGTTTTCAAAAGTTGAAGCAATTTCTAAGACCTTGAGAAAT 294

QY 304 CTCGCAAGTTTCTCCTCGCTTTGCTAAAAATTTGGTATTGAGGTGTTACTGTTTCTGATGT 363
Db 295 TCAACAGGTTTCTTCCGGCACTACTGAAAATGGGCAATTCGTGTGTCACTGTTTCAGATGT 354

QY 364 TCGAGGTTTGGTCTCAAGTGGTTCAACTGAGAGCGAGCGCGGCTCAGAAATTTCTGA 423
Db 355 TCGTGGTTTGGCGCCCAAGTGGCTTGACTGAGAGGCAAGCTGCTCTGAATTTCTCTGA 414

QY 424 AGACAAGTTTGTTCCTAAAGTTAAGATGGAGATCGTGGTTAGCAAGACCAGGTTGAGGA 483
Db 415 AGACAGTTTGTGGCAAAAGTTAAATGGAAATTTGTTGTCAGCAAGACCAGGTTGAGG 474

QY 484 TGTATAGAAAAATCAATTCAGGAGCGCAAGAACTGGAGAGATTGGAGACGGCAAGATTTT 543
Db 475 AGTCATTGCCATGATAATTGAAGAGCGCAAGAACTGGTGAATAGGTGATGGAAGATATT 534

QY 544 CTGCTGCTGCTGTTTCAGATGTAATTAAGAGTCCGCACTGTCAGCGGGGTGATAAGGCTGA 603
Db 535 CTGACTCCCATCTCCGATGTTATAGAGTTTCGCACTGGTGAACGGGGAGATAAGGCAGA 594

QY 604 GAGGATGACAGGAGGCGGATCTGACATGAGTACTTCTGCTGACTGCTGTGACCGACAA 662
Db 595 GAGGATGATGGGAGGCGCATGCTGATATGTCCTCTGCAATTATCAACTTCTTGAGTGGCAA 653

RESULT 13

AJ558383
LOCUS
DEFINITION
018 : 07 f22, mRNA sequence.
AJ558383_1
VERSION
AJ558383.1
KEYWORDS
SOURCE
ORGANISM
Antirrhinum majus (snapdragon)
Antirrhinum majus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Antirrhinaceae; Antirrhineae;
Antirrhinum.

1 (bases 1 to 732)

Zachgo,S., Stueber,K., Saedler,H., Sommer,H. and Schwarz-Sommer,Z.

Antirrhinum EST collection

Unpublished (2003)

Contact: Schwarz-Sommer Z

Molekulare Pflanzen-genetik

MPI fuer Zuechtungs-forschung

Carl-von-Linne Weg 10, D-50829, Germany.

FEATURES

Source

Location/Qualifiers

1.732
/organism="Antirrhinum majus"
/mol_type="mRNA"
/db_xref="taxon:4151"
/clone="018_1_07_f22"
/tissue_type="whole plant"
/clone_lib="Antirrhinum majus whole plant"

ORIGIN

Query Match 30.7%; Score 275; DB 9; Length 732;
Best Local Similarity 77.8%; Pred. No. 1.1e-33;
Matches 332; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 215 GTTCCTGTGATTAAAGCCCAAGCTCGCCTGACTACATTCCTGATGCTAAATTTCTACAAA 274
Db 169 GCTCCAATTGTTAGGGCTCAGATTTCTCCAGAAATCCCTACTGACTCCAAAGTTCTACAAA 228

QY 275 GTGGAAGCAATTCTCAGGCCCTGGCGAGTCTCGCAAGTTTCTCGGCTTTGCTAAAAATT 334
Db 229 GTGGAAGCTATTCTAAGACCCTGGAGGATCCAAACAGGTTTCTCGGCTTTGCTGAAAAATG 288

QY 335 GGTATTCGAGGTGTTACTGTTTCTGATGTTTCGAGGTTTGGTGGCTCAAGGTGGTTCAACT 394
Db 289 GGAATTCGGTGTCTCCTGATGTACGAGGTTTGGTTCAGGGTGGTTTGACA 348

QY 395 GAGAGCGAGGCGGCTCAGAAATTTCTGAAGACAAGTTTGTGCTAAAAGTTAAGATGGAG 454
Db 349 GAAAGACAGTTTGGCTCTGAATTTCTCAGAGACAAATTTTGTGCAAAAAGTTAAGATGGAG 408

QY 455 ATCGTGGTTAGCAAAGACCAGGTTGAGGATGTTATAGAAAAATCATTCAGGAGGCAAGA 514
Db 409 ATTGTTGTGAGCAAAGACCAGGTTGATCGGTAATAGAGAAGTATAGAGGAAGCAAGA 468

QY 515 ACTGAGAGATTGGAGACGGCAAGATTTTCTGCTGCCTGTTTCAGATGTAATAAGATC 574
Db 469 ACAGGAGAGATTGGTGTATGGGAAGATTTTGTCTCCAGTTTCTGATGTAATAAGATC 528

QY 575 CGCACTGCTGAGCGGGGTGATAAGGCTGAGAGGATGACAGGAGGCGGATCTGACATGAT 634
Db 529 CGCACTGCTGAACGAGGGGAAAGGCTGAGAGGATGTTGGGAGGACGGGCGGACATGTTT 588

QY 635 ACCTTCG 641
Db 589 TCTAATG 595

RESULT 14

AW035791

LOCUS

DEFINITION

EST281945 tomato callus, TAMU Lycopersicon esculentum cDNA clone

cLEC36D8, mRNA sequence.

AW035791

VERSION

AW035791.1

KEYWORDS

SOURCE

ORGANISM

Lycopersicon esculentum (tomato)

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 613)

Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,

Holt,I.E., Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S.,

Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and

Giovannoni,J.

Generation of ESTs from tomato callus tissue

Unpublished (1999)

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

613 bp mRNA linear EST 18-MAY-2001

AW035791

EST281945 tomato callus, TAMU Lycopersicon esculentum cDNA clone

cLEC36D8, mRNA sequence.

AW035791

VERSION

AW035791.1

KEYWORDS

SOURCE

ORGANISM

Lycopersicon esculentum (tomato)

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 613)

Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,

Holt,I.E., Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S.,

Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and

Giovannoni,J.

Generation of ESTs from tomato callus tissue

Unpublished (1999)

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

FEATURES
source

Location/Qualifiers
1. 613
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLER36D8"
/tissue_type="callus"
/dev_stage="25-40 days old"
/lab_host="XL1-Blue MRF"
/clone_lib="tomato callus, TAMU"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons
of seedlings 7-10 days post-germination were excised, cut
at both ends and placed on MS medium with no selection.
Mixed callus was harvested at 25 and 40 days and included
undifferentiated masses. Tomato Callus EST Library"

ORIGIN

Query Match 30.6%; Score 274.2; DB 9; Length 613;
Best Local Similarity 73.3%; Pred. No. 1.6e-33;
Matches 351; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 184 CGCGGTCAAGCGGTAAAGATATGCCCGCTCGTTCCTGTGATTAATGCCCAAGCTCGCC 243
Db 134 CACCTTCAACCGCTGTCAAAATGCTCTCTTCCCGATTATCAGAGCCCAAACTCTCC 193
QY 244 TGACTACATTCCTGATGCTAAATCTACAAAGTGAAGCAATCTCAGGCCCTGGCGAGT 303
Db 194 AGATTTCTGTCCTGATGCGAAGTTTACAAAGTGAAGCAATCTAAGACCTTGGAGAAT 253
QY 304 CTCGCAAGTTTCTCGGCTTTGCTAAATTTGGTATTCTGAGGTGTTACTGTTTCTGATGT 363
Db 254 TCAACAGGTTTCTTCGGCACTACTGAAATGGGCAATTCGTGGTGTCACTGTTTCGGATGT 313
QY 364 TCGAGGTTTGGTGTCAAGGTGGTTCAACTGAGAGCGGCGGCTCAGAAATTTCTGA 423
Db 314 TCGTGGTTTGGCGCCCAAGGTGGCTTGACTGAGAGCAAGCTGGCTCTGAATTTCTCTGA 373
QY 424 AGACAAAGTTTGTCTAAAGTTAAGATGGAGATCGTGGTTAGCAAAAGACCAGGTTGAGGA 483
Db 374 AGACACGTTTGTTCGCAAAAGTTAAATGGAATTTGTTGTCAGCAAAAGACCAGGTTGAAGG 433
QY 484 TGTATAGAAAATCATTGAGGAGGCAAGAACTGGAGAGATTGGAGAOGGCAAGATTTT 543
Db 434 AGTCATTGCCATGATAATTGAAGAGGCAAGAACTGGTGAATAGGTGATGGAAGATATT 493
QY 544 CTTGCTGCTGTTTTCAGATGTAATAAGATCCGCACTGGTGAGCGGGGTGATAAGGCTGA 603
Db 494 CTTGACTCCCATCTCCGATGTTATAGAGTTTCGCACTGGTGAACGGGGAGAAAGGCAGA 553
QY 604 GAGGATGACAGGAGGGCGATCTGACATGAGTACTTCTGCTTGACTGCTGTGACCAGCAA 662
Db 554 GAGGATGATGGAGGGCATGCTGATATGTCCTCTGCAATTATCAACTTCTTGAGTGGCAA 612

RESULT 15

AI773079 648 bp mRNA linear EST 18-MAY-2001
LOCUS
DEFINITION
EST254179 tomato resistant, Cornell Lycopersicon esculentum cDNA
clone cLER5E12, mRNA sequence.

AI773079
AI773079.1 GI:5271120
EST.
KEYWORDS
SOURCE

ORGANISM
Lycopersicon esculentum (tomato)
Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 648)

REFERENCE
AUTHORS

D' Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E.,
Liang, F., Opton, J., Rouning, C.M., Craven, M.B., Fujii, C.Y.,
Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Tanksley, S.D.,

TITLE
JOURNAL
COMMENT

Giovannoni, J.J. and Martin, G.B.
Generation of ESTs from Pseudomonas resistant tomato
Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES
source

Location/Qualifiers
1. 648
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="RL1-12 (35S::Pto in Rio Grande x Money Maker)"
/db_xref="taxon:4081"
/clone="cLER5E12"
/tissue_type="leaf"
/dev_stage="4-week old"
/lab_host="SOLR"
/clone_lib="tomato resistant, Cornell"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLER - Tomato Pseudomonas Resistant EST Library.
Directionally cloned cDNAs inserted into pBluescript
SK(-) at 5' end with EcoRI and 3' end with XhoI site."

ORIGIN

Query Match 30.6%; Score 274.2; DB 9; Length 648;
Best Local Similarity 75.7%; Pred. No. 1.5e-33;
Matches 339; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 184 CGCGGTCAAGCGGTAAAGATATGCCCGCTCGTTCCTGTGATTAATGCCCAAGCTCGCC 243
Db 181 CACCTTCAACCGCTGTCAAAATGCTCTCTTCCCGATTATCAGAGCCCAAACTCTCC 240
QY 244 TGACTACATTCCTGATGCTAAATTTACAAAGTGAAGCAATTTCTCAGGCCCTGGCGAGT 303
Db 241 AGATTTCTGTCCTGATGCGAAGTTTACAAAGTGAAGCAATTTCTAAGACCTTGGAGAAT 300
QY 304 CTCGCAAGTTTCTCGGCTTTGCTAAATTTGGTATTCTGAGGTGTTACTGTTTCTGATGT 363
Db 301 TCAACAGGTTTCTTCGGCACTACTGAAATTTGGCAATTCGTGGTGTCACTGTTTCGGATGT 360
QY 364 TCGAGGTTTGGTGTCTCAAGGTGGTTCAACTGAGAGCGGCGGCTCAGAAATTTTCTGA 423
Db 361 TCGTGGTTTGGCGCCCAAGGTGGCTTGACTGAGAGCAAGCTGGCTCTGAATTTCTCTGA 420
QY 424 AGACAAAGTTTGTGTCTAAAGTTAAGATGGAGATCGTGGTTAGCAAAAGACCAGGTTGAGGA 483
Db 421 AGACACGTTTGTGTGCAAAAGTTAAATTTGAAATTTGTTGTGTCAGCAAAAGACCAGGTTGAAGG 480
QY 484 TGTATAGAAAATCATTGAGGAGGCAAGAACTGGAGAGATTGGAGACGGCAAGATTTT 543
Db 481 AGTCATTGCCAGATTAATTGAGAGGCAAGAACTGGTGAATAGGTGATGGAAGATATT 540
QY 544 CTTGCTGCTGTTTTCAGATGTAATAAGAGTCCGCACTGGTGAGCGGGGTGATAAGGCTGA 603
Db 541 CTTGACTCCCATCTCCGATGTTATAAGAGTTTCGCACTGGTGAACGGNGAGAAAGGCAGA 600
QY 604 GAGGATGACAGGAGGGCGATCTGACATG 631
Db 601 GAGGATGATGGAGGGCATGCTGATATG 628

Search completed: May 28, 2004, 00:14:53
Job time : 2080.03 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 13:35:34 ; Search time 1738.01 Seconds
(without alignments)
14663.718 Million cell updates/sec

Title: US-09-756-541-15
Perfect score: 588
Sequence: 1 ATGGCGGGTCAATGACGAA.....GTGATATGCTTTCACCGTCT 588

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match: 0%
Maximum Match: 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	588	100.0	588	6	AR125591	AR125591 Sequence
2	588	100.0	591	6	AX507438	AX507438 Sequence
3	588	100.0	591	8	BT005209	BT005209 Arabidops
4	588	100.0	817	6	AR125589	AR125589 Sequence
5	588	100.0	844	8	AF095455	AF095455 Arabidops
6	285	48.5	594	6	AR125592	AR125592 Sequence
7	285	48.5	840	8	AF095454	AF095454 Ricinus c
8	285	48.5	897	6	AR125590	AR125590 Sequence
9	250.8	42.7	796	8	AY027892	AY027892 Medicago
10	245	41.7	947	8	AY442185	AY442185 Lycopersi
11	209.6	35.6	1085	8	PPI489604	AJ489604 pinus pin
12	199.8	34.0	120185	8	AC007138	AC007138 Arabidops
13	199.8	34.0	198220	8	ATCHNIV5	AL161493 Arabidops
14	199.4	33.9	902	8	AK068407	AK068407 Oryza sat
15	199.4	33.9	902	8	AK099152	AK099152 Oryza sat
16	109.8	18.7	384	1	FDGLNBPRT	X97327 F.diplosiph
17	107.4	18.3	300143	1	AE017165	AE017165 Prochloro
18	106.8	18.2	164921	8	AF022186	AF022186 Cyanidium
19	106	18.0	2493	1	ASP251822	AJ251822 Anabaena
20	106	18.0	341880	1	AP003589	AP003589 Nostoc sp
21	102	17.3	262202	1	BX572094	BX572094 Prochloro
22	101.6	17.3	191028	8	PPU38804	U38804 Porphyra pu
23	98.6	16.8	4348	1	AF079137	AF079137 Synchoco
24	98	16.7	686	1	AF017419	AF017419 Nostoc pu
25	97	16.5	383	1	SYOGLNB	M62447 Synchococc
26	95.8	16.3	339	1	PMA271089	AJ271089 Prochloro
27	95.8	16.3	11771	1	AE000674	AE000674 Aquifex a
28	94.2	16.0	349746	1	BX572099	BX572099 Prochloro
29	94	16.0	299350	1	AP005370	AP005370 Thermosyn
30	93.6	15.9	302050	1	AP006568	AP006568 Gloeobact
31	93.6	15.9	302320	1	AP005081	AP005081 Vibrio pa
32	93.2	15.9	300045	1	AE016803	AE016803 Vibrio vu
33	92.8	15.8	96109	6	AR408756	AR408756 Sequence
34	92.8	15.8	96109	6	AX067460	AX067460 Sequence
35	91.6	15.6	247950	1	AP005340	AP005340 Vibrio vu
36	89.8	15.3	11371	1	AE006236	AE006236 Pasteurel
37	89.4	15.2	351	6	AR378407	AR378407 Sequence
38	89.4	15.2	300732	1	AE016802	AE016802 Vibrio vu
39	89	15.1	1622	1	AF120107	AF120107 Synchoco
40	88.4	15.0	452	1	SSGLNBP	X97496 Synchococyst
41	88.4	15.0	130001	1	D90915	D90915 Synchococyst
42	86.8	14.8	298950	1	AP004597	AP004597 Oceanobac
43	86.2	14.7	248650	1	AP005341	AP005341 Vibrio vu
44	85.8	14.6	375	6	AR318761	AR318761 Sequence
45	85.4	14.5	405	6	AR376903	AR376903 Sequence

ALIGNMENTS

RESULT 1
AR125591
LOCUS AR125591 588 bp DNA
DEFINITION Sequence 15 from patent US 6177275.
ACCESSION AR125591
VERSION AR125591.1 GI:14111653
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 588)
AUTHORS Coruzzi,G.M., Lam,H.-M. and Hsieh,M.-H.
TITLE Plant nitrogen regulatory P-II genes
JOURNAL Patent: US 6177275-A 15 23-JAN-2001;
FEATURES Location/Qualifiers

source		..588		/organism="unknown"		/mol_type="unassigned DNA"													
ORIGIN		Query Match		100.0%; Score 588; DB 6; Length 588;		Best Local Similarity		100.0%; Pred. No. 2.8e-137;		Matches 588; Conservative		0; Mismatches		0; Indels		0; Gaps		0;	
QY		1		ATGGCGGCGTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTTCTATTTCTGATCGA		60													
Db		1		ATGGCGGCGTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTTCTATTTCTGATCGA		60													
QY		61		AAGAACATTGCTTTCTCTGATTGCAATTCGATTGTTCTGGATTTCAGACATTCCTCGACCA		120													
Db		61		AAGAACATTGCTTTCTCTGATTGCAATTCGATTGTTCTGGATTTCAGACATTCCTCGACCA		120													
QY		121		TCCTGCTCGATTGTCACAAAGTCACCGAGTAATAACAGTCGTCGTTTACCTGTCGTT		180													
Db		121		TCCTGCTCGATTGTCACAAAGTCACCGAGTAATAACAGTCGTCGTTTACCTGTCGTT		180													
QY		181		AGTGCCCAATATCTTCTGATTATATTCAGACTCGAAATTTTACAAGTGGAGCAATT		240													
Db		181		AGTGCCCAATATCTTCTGATTATATTCAGACTCGAAATTTTACAAGTGGAGCAATT		240													
QY		241		GTCAGACCATGGAGAAATCCAGCAAGTTTCATCGGCTTTACTGAAATCGGATTCGAGGT		300													
Db		241		GTCAGACCATGGAGAAATCCAGCAAGTTTCATCGGCTTTACTGAAATCGGATTCGAGGT		300													
QY		301		GTTACTGTTTCTGATGTGAGAGGGTTTGGTGCAAGGAGTTCTACCGAGAGACACGGT		360													
Db		301		GTTACTGTTTCTGATGTGAGAGGGTTTGGTGCAAGGAGTTCTACCGAGAGACACGGT		360													
QY		361		GGCTCTGAGTTCTCGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAG		420													
Db		361		GGCTCTGAGTTCTCGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAG		420													
QY		421		AAAGACCAAGTGGAAATCTGTAATCAACACAAATTAATGAAGGAGCAAGGACAGAGATT		480													
Db		421		AAAGACCAAGTGGAAATCTGTAATCAACACAAATTAATGAAGGAGCAAGGACAGAGATT		480													
QY		481		GGTGATGGCAAGATTTTGTTCCTGTGTGATGTCATAGAGTTAGGACAGGTGAG		540													
Db		481		GGTGATGGCAAGATTTTGTTCCTGTGTGATGTCATAGAGTTAGGACAGGTGAG		540													
QY		541		CGTGGGGAGAAAGCAGAGAAGATGACTGGTGATATGCTTTCACCGTCT		588													
Db		541		CGTGGGGAGAAAGCAGAGAAGATGACTGGTGATATGCTTTCACCGTCT		588													
RESULT 2		AX507438		LOCUS		AX507438		Sequence 2133 from Patent WO0216655.		591 bp		DNA		linear		PAT 27-SEP-2002			
DEFINITION		Arabidopsis thaliana		Arabidopsis thaliana (thale cress)		Arabidopsis thaliana		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;											
ACCESSION		AX507438		VERSION		AX507438.1		GI:23388675											
KEYWORDS																			
SOURCE																			
REFERENCE																			
AUTHORS																			
TITLE																			
JOURNAL																			
FEATURES		source		Location/Qualifiers		1..591		/organism="Arabidopsis thaliana"		/mol_type="unassigned DNA"									
TITLE		Arabidopsis ORF clones		Arabidopsis ORF clones		Unpublished		2 (bases 1 to 591)											
REFERENCE																			
AUTHORS																			
TITLE																			
JOURNAL																			
FEATURES		source		Location/Qualifiers		1..591													

Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Hsuan,V.W., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.

Direct Submission
Submitted (14-MAR-2003) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEN (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Chan, M.M., Chang,C.M., Dale,J.M., Hsuan,V.W., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera, C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Davis,R.W., Theologis,A., and Ecker,J.R.

Kim,C.J. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES
source
Location/Qualifiers
1..591
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="4"
/clone="U23463"
/ecotype="Columbia"
/note="This clone is in pUNI 51"
1..591
/note="P II nitrogen sensing protein GLB 1"
/codon_start=1
/product="At4g01900"
/protein_id="AA063273.1"
/db_xref="GI:28950699"
/translation="MAASMTKPISTISLGFVSDRKNIAFSDCISICSGFRHSRPSCLD LVTKSPNNSRVLVPVSAQISSDYIPDSKFYKVEAIVRWIRQQVSSALLKIGIRGVT VSDVRFGAQQGSTERHGGSEFSEDFKFAKVKMEIVVKKDQVESVINTIIEGARTGEI GDGKIFVLPVSDVIRVIRVTGERGEKAEXMTGDMLSPS"

CDS
100.0%; Score 588; DB 8; Length 591;
Best Local Similarity 100.0%; Pred. No. 2.8e-137;
Matches 588; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
1 ATGGCGGCGTCAATGACGAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGA 60
1 ATGGCGGCGTCAATGACGAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGA 60
61 AAGAACAATTGCTTCTCTGATTGCAATTCGATTGTTGTTCTGGATTCCAGCATTCGCCACCA 120
61 AAGAACAATTGCTTCTCTGATTGCAATTCGATTGTTGTTCTGGATTCCAGCATTCGCCACCA 120
121 TCTTGCCCTCGATTGGTCACAAAGTCACCGAGTAATAACAGTCGTGTTTACCTGTCGTT 180
121 TCTTGCCCTCGATTGGTCACAAAGTCACCGAGTAATAACAGTCGTGTTTACCTGTCGTT 180
181 AGTGCCCAAAATATCTTCTGATTATATTCAGACTCGAAATTTTACAAGGTGGAAGCAATT 240
181 AGTGCCCAAAATATCTTCTGATTATATTCAGACTCGAAATTTTACAAGGTGGAAGCAATT 240
241 GTCAGACCATGGAGAATCCAGCAAGTTTCATCGGCTTTTACGAAATCGGGATTCGAGGT 300

Db 241 GTCAGACCATGGAGAATCCAGCAAGTTTTCATCGGCTTTTACTGAAAAATCGGGATTCGAGGT 300

QY 301 GTTACTGTTTCTGATGTCAGAGGGTTTGGTGCAACAAGGAGGTTCACCGAGAGACACCGGT 360

Db 301 GTTACTGTTTCTGATGTCAGAGGGTTTGGTGCAACAAGGAGGTTCACCGAGAGACACCGGT 360

QY 361 GGCTCTGAGTTCTCGGAAGCAAAATTTGTTGCTAAAGTTTAAGATGGAATCGTTGTTAAG 420

Db 361 GGCTCTGAGTTCTCGGAAGCAAAATTTGTTGCTAAAGTTTAAGATGGAATCGTTGTTAAG 420

QY 421 AAAGACCAAGTGGRAATCTGTAATCAACACAAATAATTGAAGGAGCAAGGACAGGAGATT 480

Db 421 AAAGACCAAGTGGRAATCTGTAATCAACACAAATAATTGAAGGAGCAAGGACAGGAGATT 480

QY 481 GGTGATGGCAAGATTTTGTGTTTTCCTGTCAGATGTCATTAAGAGTTAGGACAGGTGAG 540

Db 481 GGTGATGGCAAGATTTTGTGTTTTCCTGTCAGATGTCATTAAGAGTTAGGACAGGTGAG 540

QY 541 CGTGGGGAGAAAAGCAGAGAGATGACTGGTGATATGCTTTCACCGTCT 588

Db 541 CGTGGGGAGAAAAGCAGAGAGATGACTGGTGATATGCTTTCACCGTCT 588

RESULT 4
AR125589
LOCUS AR125589 817 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 13 from patent US 6177275.
ACCESSION AR125589
VERSION AR125589.1 GI:14111651
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 817)
AUTHORS Coruzzi,G.M., Lam,H.-M. and Hsieh,M.-H.
TITLE Plant nitrogen regulatory P-PII genes
JOURNAL Patent: US 6177275-A 13 23-JAN-2001;
FEATURES
source
1..817
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 588; DB 6; Length 817;
Best Local Similarity 100.0%; Pred. No. 2.7e-137;
Matches 588; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGCGTCAATGACGAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGA 60
Db 33 ATGGCGGCGTCAATGACGAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGA 92

QY 61 AAGAACAATTGCTTCTCTGATTGCAATTCGATTGTTGTTCTGGATTCCAGCATTCGCCACCA 120
Db 93 AAGAACAATTGCTTCTCTGATTGCAATTCGATTGTTGTTCTGGATTCCAGCATTCGCCACCA 152

QY 121 TCTTGCCCTCGATTGGTCACAAAGTCACCGAGTAATAACAGTCGTGTTTACCTGTCGTT 180
Db 153 TCTTGCCCTCGATTGGTCACAAAGTCACCGAGTAATAACAGTCGTGTTTACCTGTCGTT 212

QY 181 AGTGCCCAAAATATCTTCTGATTATATTCAGACTCGAAATTTTACAAGGTGGAAGCAATT 240
Db 213 AGTGCCCAAAATATCTTCTGATTATATTCAGACTCGAAATTTTACAAGGTGGAAGCAATT 272

QY 241 GTCAGACCATGGAGAATCCAGCAAGTTTCATCGGCTTTTACGAAATCGGGATTCGAGGT 300
Db 273 GTCAGACCATGGAGAATCCAGCAAGTTTCATCGGCTTTTACTGAAATCGGGATTCGAGGT 332

QY 301 GTTACTGTTTCTGATGTCAGAGGGTTTGGTGCAACAAGGAGGTTCACCGAGAGACACCGGT 360
Db 333 GTTACTGTTTCTGATGTCAGAGGGTTTGGTGCAACAAGGAGGTTCACCGAGAGACACCGGT 392

QY 361 GGCTCTGAGTTCTCGGAAGCAAAATTTGTTGCTAAAGTTTAAGATGGAATCGTTGTTAAG 420

Db 244 CCTGGCGAGTCTCGCAAGTTTCCTCGGCTTTGCTAAAAATGGTAATTCGAGGTGTACT 303
Qy 307 GTTCTGATGTGAGAGGGTTTGGTGCACAGGAGTTCTACCGAGAGACACGGTGGCTCT 366
Db 304 GTTCTGATGTGAGAGTTTGGTGTCTCAAGTGGTTCAACTGAGAGGACGGCGGCTCA 363
Qy 367 GAGTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAGAAAGAC 426
Db 364 GAAATTTCTGAAGACAAAGTTTGTGCTAAAGTTAAGATGGAGATCGTGGTTAGCAAAGAC 423
Qy 427 CAAGTGGAAATCTGTAATCAACACAATAATTGAAGGAGCAAGGACAGGAGATGGTGAT 486
Db 424 CAGGTTGAGGATGTTATAGAAAAAATCAATTGAGGAGGCAAGAACTGGAGAGATGGAGAC 483
Qy 487 GGCAAGATTTTGTGTTGCTGCTGTGTCAGATGTCATAAGAGTTAGGACAGGTGAGCGTGG 546
Db 484 GGCAAGATTTTCTGCTGCTGCTGTTTCAGATGTAATAAGAGTCCGCACCTGGTGAGCGGGT 543
Qy 547 GAGAAAGCAGAGAGATGACTGGTG 571
Db 544 GATAAGGCTGAGAGGATGACAGGAG 568

RESULT 7
AF095454 840 bp mRNA linear PLN 19-NOV-1998
LOCUS Ricinus communis PII protein mRNA, partial cds.
DEFINITION AF095454
ACCESSION AF095454.1 GI:3885940
KEYWORDS Ricinus communis (castor bean)
SOURCE Ricinus communis
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Malpighiales; Euphorbiaceae; Acalyphoideae; Acalyphaceae; Ricinus.

REFERENCE 1 (bases 1 to 840)
AUTHORS Hsieh,M.H., Lam,H.M., van de Loo,F.J. and Coruzzi,G.
TITLE A PII-like protein in Arabidopsis: putative role in nitrogen sensing
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (23), 13965-13970 (1998)
MEDLINE 99030678
PUBMED 9811909
REFERENCE 2 (bases 1 to 840)
AUTHORS Hsieh,M.H., Lam,H.M., van de Loo,F.J. and Coruzzi,G.M.
TITLE Direct Submission
JOURNAL Submitted (29-SEP-1998) Biology, New York University, 100 Washington Square East, New York, NY 10003, USA

FEATURES
source Location/Qualifiers
1..840 /organism="Ricinus communis"
/mol_type="mRNA"
/db_xref="taxon:3988"
CDS <1..595
/function="putative role in nitrogen sensing"
/note="GLB1"
/codon_start=2
/product="PII protein"
/protein_id="AAC78332.1"
/db_xref="GI:3885941"
/translation="HEATAKLGLLTPLHNSNNIKKEFPVDFSLFCELPRLHRSRPSHENT
AVKRVYAPVPVINAQSPDYPDAKFYKVEAILRPMRVSOVSSALLKIGIRGVTVS
DVRGFGAQQGSTERQGSSEFSEDFVAKVMEIVVSKQVEDVIEKIIERARTGEIGD
GKIFLLPVSVDIVRTGERGDKAERMTGGRSDMSTA"

ORIGIN
Query Match 48.5%; Score 285; DB 8; Length 840;
Best Local Similarity 70.3%; Pred. No. 4.7e-61;
Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

Qy 7 GCGTCAATGACGAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGAAAGAAC 66
Db 5 GAGGCTACTGCGAAACTGGGCTTGCTCACTCTCTTCAITTAATAACATCAAGAAGAA 64

Qy 67 ATTGCTTTCTCTGATTGCAATTCGATTGTTCTGGATTTCAGACATTCCTCCGACCATCTTGC 126
Db 65 TTCCCTGTTTTGATTTCAGTTTGTGTTGTCAGAGCTTAGACATTTCTCGGTTTCTCAC 124
Qy 127 CTCGATTGGTCACAAAGTCACCGAGTAAATAACAGTCGTGTTTTACCTGTCTGTTAGTSCC 186
Db 125 TTAAACACCGCGGTCAAGCGCGTAAGATATGCC---CCCCTCGTTCCTGTGATTAATGCC 181
Qy 187 CAAATATCTTCTGATTATATTCAGACTCGAAATTTTACAAGGTGGAAGCAATTGTCTAGA 246
Db 182 CAAAGCTCGCTGACTACATTCCTGATGCTAAATTTCTACAAAGTGAAGCAATTTCTCAGG 241
Qy 247 CCATGGAGAAATCCAGCAAGTTTTCATCGGCTTTACTGAAATTCGGGATTCGAGGTGTACT 306
Db 242 CCCTGGGAGTCTCGCAAGTTTCTCGGCTTTGCTAAAAAATTGGTATTTCGAGGTGTACT 301
Qy 307 GTTCTGATGTGAGAGGGTTTGGTGCACAAAGGAGGTCTACCGAGAGACACCGTGGCTCT 366
Db 302 GTTCTGATGTTCGAGGTTTGGTGCTCAAGGTGGTTCAACTGAGAGGCGGCGGCTCA 361
Qy 367 GAGTCTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAGAAAGAC 426
Db 362 GAATTTTCTGAAGACAAAGTTTGTGCTAAAGTTAAGATGAGATCGTGGTTAGCAAAGAC 421
Qy 427 CAAGTGGAAATCTGTAATCAACACAATAATTGAAGGAGCAAGGACAGGAGATTTGGTGAT 486
Db 422 CAGGTTGAGGATGTTATAGAAAAAATCAATTGAGGAGGCAAGAACTGGAGAGATTTGAGAC 481
Qy 487 GGCAAGATTTTGTGTTTGGTGTGTGATGTCATAGAGTTAGGACAGGTGAGCGTGGG 546
Db 482 GGCAAGATTTTCTTGCTGCTGCTGTTTCAGATGTAATAAGAGTCCGCACTGGTGAGCGGGT 541
Qy 547 GAGAAAGCAGAGAGATGACTGGTG 571
Db 542 GATAAGGCTGAGAGGATGACAGGAG 566

RESULT 8
AR125590 897 bp DNA linear PAT 16-MAY-2001
LOCUS AR125590
DEFINITION Sequence 14 from patent US 6177275.
ACCESSION AR125590
VERSION AR125590.1 GI:14111652
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 897)
AUTHORS Coruzzi,G.M., Lam,H.-M. and Hsieh,M.-H.
TITLE Plant nitrogen regulatory P-PII genes
JOURNAL Patent: US 6177275-A 14 23-JAN-2001;
FEATURES Location/Qualifiers
source 1..897
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 48.5%; Score 285; DB 6; Length 897;
Best Local Similarity 70.3%; Pred. No. 4.7e-61;
Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

Qy 7 GCGTCAATGACGAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGAAAGAAC 66
Db 56 GAGGCTACTGCGAAACTGGGCTTGCTCACTCTCTTCAITTAATAACATCAAGAAGAA 115
Qy 67 ATTGCTTTCTCTGATTGCAATTCGATTGTTCTGGATTTCAGACATTCCTCCGACCATCTTGC 126
Db 116 TTCCCTGTTTTGATTTCAGTTTGTGTTGTCAGAGCTTAGACATTTCTCGGTTTCTCAC 175
Qy 127 CTCGATTGGTCACAAAGTCACCGAGTAAATAACAGTGTGTTTTTACCTGTCTGTAGTCC 186
Db 176 TTTAACACCGCGGTCAAGCGCGTAAGATATGCC---CCCCTCGTTCCTGTGATTAATGCC 232

		Matches	309;	Conservative	0;	Mismatches	97;	Indels	0;	Gaps	0;
QY	187	CAATATCTTCTGATTATATCCAGACTCGAAATTTTACAAAGTGGAGCAATTTGTCTAGA	246								
Db	233	CAAAAGCTGCCTGACTACATTCCTGATGCTAAATTTCTACAAAGTGGAGCAATTTCTCAGG	292								
QY	247	CCATGGAGAAATCCAGCAAGTTTTCATCGGCTTTTACTGAAATTCGGGATTCGAGGTGTTACT	306								
Db	293	CCCTGGCGAGTCTCGCAAGTTTCTCGGCTTTGCTAAATAATGGTATTCGAGGTGTTACT	352								
QY	307	CTTCTGATGTGAGAGGGTTTGGTGCAAGGAGGTCTACCGAGAGACACGGTGGCTCT	366								
Db	353	GTTCCTGATGTTCGAGGTTTTGGTGCTCAAGGTGGTTCAACTGAGAGCGCAGGGCGGCTCA	412								
QY	367	GAGTTCCTGGAAGACAAATTTGTTGCTTAAAGTTAAGATGGAATCGTTGTTAAGAAAGAC	426								
Db	413	GAATTTTCTGAAGACAAAGTTTGTGTTGCTTAAAGTTAAGATGGAGATCGTGTAGCAAGAC	472								
QY	427	CAAGTGGATCTGTAATCAACACAATAATCAAGGAGCAAGGACAGGAGATTTGGTGAT	486								
Db	473	CAGGTTGAGGATGTTATAGAAAATAATCAATGAGGAGGCAAGAACTGGAGAGATTTGGAGAC	532								
QY	487	GGCAAGATTTTGTGCTGCTGTCAGATGTCATAAGAGTTAGGACAGGTGAGCGTGGG	546								
Db	533	GGCAAGATTTTCTGCTGCTGCTGTTTCAGATGTAATAAGAGTCGCACTGCTGAGCGGGGT	592								
QY	547	GAGAAAGCAGAGAGATGACTGGTG	571								
Db	593	GATAAGGCTGAGAGATGACAGGAG	617								

RESULT 9	AY027892	796 bp	mRNA	linear	PLN 12-MAR-2001
LOCUS	Medicago sativa PII protein (GLN3)	mRNA, complete cds.			
DEFINITION	AY027892				
ACCESSION	AY027892.1	GI:13277514			
VERSION					
KEYWORDS					
SOURCE	Medicago sativa				
ORGANISM	Medicago sativa				
REFERENCE	1 (bases 1 to 796)				
AUTHORS	Garcia-Ibáñez, D. and Sengupta-Gopalan, C.				
TITLE	Characterization of PII (GLN3) in alfalfa				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 796)				
AUTHORS	Garcia-Ibáñez, D. and Sengupta-Gopalan, C.				
TITLE	Direct Submission				
JOURNAL	Submitted (22-FEB-2001) Molecular Biology Program, New Mexico State University, Corner of Knox and College, Las Cruces, NM 88003, USA				
FEATURES	Location/Qualifiers				
source	1..796				
	/organism="Medicago sativa"				
	/mol_type="mRNA"				
	/db_xref="taxon:3879"				
gene	1..796				
	/gene="GLN3"				
CDS	1..585				
	/gene="GLN3"				
	/codon_start=1				
	/product="PII protein"				
	/protein_id="AAK16221.1"				
	/db_xref="GI:13277515"				
	/translation="MTLIAKPNVFNGLNFINETQIPFSSFSVIRKRFQDSHNRNVL				
	RSGNASVLPKIRAQQLPDYVPESKFKYKVEAILRPWRIQVSSGLKRMGIRGVTVDV				
	KGFGAQQGSKERQGGSEFSDNFVAKVKMEIVVRKQDQVEAVINKIMETARTGEIGDGK				
	IFLIPVSDVIRIRTGERGEQAERMAGGLTDALYV"				
ORIGIN					
Query Match	42.7%; Score 250.8; DB 8; Length 796;				
Best Local Similarity	76.1%; Pred. No. 1.9e-52;				

RESULT 10	AY442185	947 bp	mRNA	linear	PLN 15-NOV-2003
LOCUS	Lycopersicon esculentum PII-like protein (GLB1)	mRNA, complete cds.			
DEFINITION	AY442185				
ACCESSION	AY442185.1	GI:38231569			
VERSION					
KEYWORDS					
SOURCE	Lycopersicon esculentum (tomato)				
ORGANISM	Lycopersicon esculentum				
REFERENCE	1 (bases 1 to 947)				
AUTHORS	Walch-Liu, P., Roemheld, V. and von Wiren, N.				
TITLE	A PII-like protein in tomato				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 947)				
AUTHORS	Walch-Liu, P., Roemheld, V. and von Wiren, N.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-OCT-2003) Institute of Plant Nutrition, Hohenheim University, Fruwirthstr.20, Stuttgart 70593, Germany				
FEATURES	Location/Qualifiers				
source	1..947				
	/organism="Lycopersicon esculentum"				
	/mol_type="mRNA"				
	/db_xref="taxon:4081"				
gene	1..947				
	/gene="GLB1"				
	/note="LeGLB1"				
CDS	53..652				
	/gene="GLB1"				
	/note="similar to nitrogen regulatory protein PII"				
	/codon_start=1				
	/product="PII-like protein"				
	/protein_id="AAK14689.1"				
	/db_xref="GI:38231570"				
	/translation="MASPSLSKSNFSLHSFSSPSLSQPHFTSITVVQKPFPSQLT				
	FKRCQNAFSPFPIRAQNSPDFVPDAKFYKVEAILRPWRIQVSSALLKMGIRGVTVDV				
	VRGFGAQQGLTERQAGSEFSDTFVAKVKMEIVVSKQVQVEGVIAKIIEEARTGEIGDG				
	KIFLTPISDVIRVTGERGEKAERMVGGHADMSALSTS"				
ORIGIN					

RESULT 10	AY442185	947 bp	mRNA	linear	PLN 15-NOV-2003
LOCUS	Lycopersicon esculentum PII-like protein (GLB1) mRNA, complete cds.				
DEFINITION	AY442185				
ACCESSION	AY442185.1	GI:38231569			
VERSION					
KEYWORDS					
SOURCE	Lycopersicon esculentum (tomato)				
ORGANISM	Lycopersicon esculentum				
REFERENCE	1 (bases 1 to 947)				
AUTHORS	Walch-Liu, P., Roemheld, V. and von Wieren, N.				
TITLE	A PII-like protein in tomato				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 947)				
AUTHORS	Walch-Liu, P., Roemheld, V. and von Wieren, N.				
TITLE	Direct Submission				
JOURNAL	Submitted {17-OCT-2003} Institute of Plant Nutrition, Hohenheim University, Fruwirthstr.20, Stuttgart 70593, Germany				
FEATURES	Location/Qualifiers				
source	1..947				
	/organism="Lycopersicon esculentum"				
	/mol_type="mRNA"				
	/db_xref="taxon:4081"				
gene	1..947				
	/gene="GLB1"				
	/note="LeGLB1"				
CDS	53..652				
	/gene="GLB1"				
	/note="similar to nitrogen regulatory protein PII"				
	/codon_start=1				
	/product="PII-like protein"				
	/protein_id="AAR14689.1"				
	/db_xref="GI:38231570"				
	/translation="MASPSLSKSNFSLHFSFSSPSLSQFPHFTSITVVQKFPSPQLT				
	FKRCQNAFPIIRAQNSPDFVDAKFYKVEAILRPWRIQQVSSALLKMGIRGVTVD				
	VRFGAQGGGLTERQAGSEFSEDTFVAKVKMEIVVSKDQVEVIKIEEARTGEIGDG				
	KIFLTPISDVIRVIRTGERGEKAERMGGHADMSALSTS"				
ORIGIN					

JOURNAL Submitted (01-APR-1999) Lita Annenberg Hazen Genome Sequencing Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724

REMARK Arabidopsis thaliana BAC T7B11 from chromosome IV near 9.5 cM

COMMENT The T7 end of T7B11 is oriented toward the NOR and overlaps with T15B16, GenBank accession number AF104919. The SP6 end of T7B11 is oriented toward the centromere and overlaps with T10M13, GenBank accession number AF001308. There are no discrepancies found in these independently finished overlaps. The region from position 110064 to 110253 exists as single stranded and single chemistry. Please see the feature below for details.

FEATURES

source location/Qualifiers

1. .120185

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/cultivar="Columbia"

/db_xref="taxon:3702"

/chromosome="IV"

/map="near 10 cM"

misc_feature 1. .14923

/note="overlap with T15B16, GenBank accession number AF104919, from position 1 to 14923"

50. .1880

/gene="T7B11.1"

/note="encodes hypothetical protein; identical to T15B16.9, GenBank accession number AF104919; gene model last edited on 3 Mar 99"

/evidence=not_experimental

join(50. .556,1098. .1259,1340. .1561,1668. .1880)

/gene="T7B11.1"

/note="identical to T15B16.9, GenBank accession number AF104919; similar to T7B11.3; functional catalog ID=99"

/codon_start=1

/evidence=not_experimental

/product="hypothetical protein"

/protein_id="AAD22637.1"

/db_xref="GI:4558544"

/translation="MAQQQQLHQRPFSPPSSRSYSSISNRPIFLSRNGLLVLLAL FLLEGVFLPWPSPLLLPFNKVSPPSYASSLSPHAKSEWRNYTLAQAAKFVATNGTVI VCAVSSPFLPFLNNWLISVSRQKHQKLVIAEDYITLYKVKWPGHVLIPPALDS DPATYSGSGFFNTARRPQHLLQILELGYNVMYNDVMWLQDPFPQYLEGSHDAYFT KMPQIKPLNHSDDLPAPONGRTYICSCMIYLRPTNGAKLMLKWSSELQSAWSES IIRPKANDQPAFNALNKAHQVDLYLLSQVAFPTGGLYFNDAAWVKETGKHVIVHNN YIIGYDRKMRFPQDYGLWLVDHLESPGLKQ"

complement(2647. .4650)

/gene="T7B11.2"

/note="encodes putative CHP-rich zinc finger protein; identical to T15B16.6, GenBank accession number AF104919; gene model last edited on 3 Mar 99"

/evidence=not_experimental

complement(2647. .4650)

/genes="T7B11.2"

/note="identical to T15B16.6, GenBank accession number AF104919; similar to A. thaliana CHP-rich proteins encodes by T10M13, GenBank accession number AF001308; functional catalog ID=98"

/codon_start=1

/evidence=not_experimental

/product="putative CHP-rich zinc finger protein"

/protein_id="AAD22638.1"

/db_xref="GI:4558545"

/translation="MDSGVSVPLIHEHLMWPCNDLRRGDCCELEAISDGYCTTCD FFVHKKCGDEASECIENPFNSHNPRLGFLSRLQROHMKVRSCLGKNIGDLFYR CEI CDPMDLHCAKYPPEVIDI PEMHSHKMLLKDVFQCDKACGKIGYGFPECH ECDSKPHVDCVRYSSSEEVHSHLPLKLLMGQPPDYSDGKCRCLGCRKID DKLFFYHCCSNFTLDMRCVLPFPKSVLDLKTHDQNLNLLPRLIFFTCNACGLNGDRS PYACFQCDFLIHKOCFGLPRLININRHDRISRTSLIGVNSVCGVCQKVDWTGCGY SCQRCSTYIAHSCATREDVWNGKELEGVPEIEIDIEPYVIDDNTIQFSHKEHYR LNVRGLMCEVNRKCNACNHPISPOQSYFGMCDDFILHONCAGFPRRKHVHLNRLAL VTSEVNI FGCSACHKIFNGFRIYEHEDTKLDVLCGSFSEPFHPSPHPHPLYCISPEDD EVCSGCNSRYHVLRCIEDNCGFILDFGCATFPQVVKHRIEDQPLSLCYGEKASGKYW CDICEKETNPNTFTCKOHRASLHTWCVLGDFMGLIPKSTIELWNISYEVVLLNNSIS RPICRHCKSHCIPPIILKKIGTSDPYFCSLDLCIESFKRLWRAK"

gene

8080. .10112

/gene="T7B11.3"

/note="encodes hypothetical protein; identical to T15B16.8, GenBank accession number AF104919; gene model last edited on 3 Mar 99"

/evidence=not_experimental

join(8080. .8568,9318. .9479,9557. .9778,9900. .10112)

/gene="T7B11.3"

/note="identical to T15B16.8, GenBank accession number AF104919; similar to T7B11.1; functional catalog ID=99"

/codon_start=1

/evidence=not_experimental

/product="hypothetical protein"

/protein_id="AAD22639.1"

/db_xref="GI:4558546"

/translation="MEQKQHILKQSTFSSPSSYSSISDRPISLLSRNGLLLLLLALV LLLGVLLPWPSPFLFPNRLSSSLSPSQSKWRDYTLAQAAKFVAKNGTVIVCAVSS PFLPFLNNWLISVSRQKHQKLVIAEDYITLYKVKWPGHVLIPPALDSKTAFSF GSQGFNFNTARRPQHLLQILELGYNVMYNDVMWLQDPFYLEGSHDAYFTDDMPQI KPLNHSDDLPHDRNGETIYICSCMIYLRPTNGAKLMLKWSSELQSAWSESIRFKAN DQPAFNALNKAHQVDLYLLSQVAFPTGGLYFNKNEAWVQETGKHVIVHNNYIIGYD RKMRFQDYGLWLVDHLESPGLKLE"

10826. .12762

/gene="T7B11.4"

/note="encodes hypothetical protein; identical to T15B16.7, GenBank accession number AF104919; gene model last edited on 3 Mar 99"

/evidence=not_experimental

join(10826. .10842,10929. .11371,11559. .11770,11843. .12130, 12260. .12433,12526. .12762)

/gene="T7B11.4"

/note="identical to T15B16.7, GenBank accession number AF104919; functional catalog ID=99"

/codon_start=1

/evidence=not_experimental

/product="hypothetical protein"

/protein_id="AAD22640.1"

/db_xref="GI:4558547"

/translation="MGTRHCVGLGYEMFNFWSSSVFVWSSSETLFKIVITRSLC ILGVHTLLTFKVLTFLLXIELQVFCVDCVVAETGIVVNIPTTKAQDGRPAGESGSK LRDEYILRLNGLLFDKAYRVDGKGKMMKXGPKGLGYGTAGADDYNDVGREGE ESAKLNAIQERTMRHIQIVDDHERLTKLLESEKKLEIKGNELAKPQVHNGTERMKL SEDLEORKEELHEKIIIRLERQIDQVQAIETEVEQLKGQLNVMMKMSDGDQAQVKEV DIIFKDLVEKBAELADLNKFNQTLILRERTNDELQEARKELVNMGELVRKPFVDM QQKYCQDVEDRAVEVLQWHEYINDPDWHPYKRKLENDREVEVIDDRDEKLRELK ADLGDGPNVATKALLEINEYNPSCRYITTELMNFEDKRAITLEEGVTCILDQWEKAK RKRGM"

13384. .14244

/gene="T7B11.5"

/note="encodes RNaseP-associated protein; identical to T15B16.20, GenBank accession number AF104919; gene model last edited on 3 Mar 99"

/evidence=not_experimental

join(13384. .13454,13566. .13613,13697. .13757,13842. .13940, 14020. .14244)

/gene="T7B11.5"

/note="identical to T15B16.20, GenBank accession number AF104919; functional catalog ID=98"

/codon_start=1

/evidence=not_experimental

/product="putative RNaseP-associated protein"

/protein_id="AAD22641.1"

/db_xref="GI:4558548"

/translation="MRGVTPINPQKKIADSNPVKESDYEGERLTHLLDLIQRGIET SKLSNVNSLPKIKQIAIGINEVTVRLERMNPNTSDQQNPKQLQVILVADCK PRMLTKHIPNLAASRNVPVLYVRDNKRASRLGELVKLTALAIIGIKARGNDLNLK QILPRDS"

complement(14442. .19647)

/gene="T7B11.6"

/note="encodes putative SecA-type chloroplast protein transport factor; gene model last edited on 26 Mar 99"

/evidence=not_experimental

complement(join(14442. .14579,14663. .14902,14991. .15108,

CDS

10826. .12762

/gene="T7B11.4"

/note="encodes hypothetical protein; identical to T15B16.7, GenBank accession number AF104919; gene model last edited on 3 Mar 99"

/evidence=not_experimental

join(10826. .10842,10929. .11371,11559. .11770,11843. .12130, 12260. .12433,12526. .12762)

/gene="T7B11.4"

/note="identical to T15B16.7, GenBank accession number AF104919; functional catalog ID=99"

/codon_start=1

/evidence=not_experimental

/product="hypothetical protein"

/protein_id="AAD22640.1"

/db_xref="GI:4558547"

/translation="MGTRHCVGLGYEMFNFWSSSVFVWSSSETLFKIVITRSLC ILGVHTLLTFKVLTFLLXIELQVFCVDCVVAETGIVVNIPTTKAQDGRPAGESGSK LRDEYILRLNGLLFDKAYRVDGKGKMMKXGPKGLGYGTAGADDYNDVGREGE ESAKLNAIQERTMRHIQIVDDHERLTKLLESEKKLEIKGNELAKPQVHNGTERMKL SEDLEORKEELHEKIIIRLERQIDQVQAIETEVEQLKGQLNVMMKMSDGDQAQVKEV DIIFKDLVEKBAELADLNKFNQTLILRERTNDELQEARKELVNMGELVRKPFVDM QQKYCQDVEDRAVEVLQWHEYINDPDWHPYKRKLENDREVEVIDDRDEKLRELK ADLGDGPNVATKALLEINEYNPSCRYITTELMNFEDKRAITLEEGVTCILDQWEKAK RKRGM"

13384. .14244

/gene="T7B11.5"

/note="encodes RNaseP-associated protein; identical to T15B16.20, GenBank accession number AF104919; gene model last edited on 3 Mar 99"

/evidence=not_experimental

join(13384. .13454,13566. .13613,13697. .13757,13842. .13940, 14020. .14244)

/gene="T7B11.5"

/note="identical to T15B16.20, GenBank accession number AF104919; functional catalog ID=98"

/codon_start=1

/evidence=not_experimental

/product="putative RNaseP-associated protein"

/protein_id="AAD22641.1"

/db_xref="GI:4558548"

/translation="MRGVTPINPQKKIADSNPVKESDYEGERLTHLLDLIQRGIET SKLSNVNSLPKIKQIAIGINEVTVRLERMNPNTSDQQNPKQLQVILVADCK PRMLTKHIPNLAASRNVPVLYVRDNKRASRLGELVKLTALAIIGIKARGNDLNLK QILPRDS"

complement(14442. .19647)

/gene="T7B11.6"

/note="encodes putative SecA-type chloroplast protein transport factor; gene model last edited on 26 Mar 99"

/evidence=not_experimental

complement(join(14442. .14579,14663. .14902,14991. .15108,

gene

13384. .14244

/gene="T7B11.5"

/note="encodes RNaseP-associated protein; identical to T15B16.20, GenBank accession number AF104919; gene model last edited on 3 Mar 99"

/evidence=not_experimental

join(13384. .13454,13566. .13613,13697. .13757,13842. .13940, 14020. .14244)

/gene="T7B11.5"

/note="identical to T15B16.20, GenBank accession number AF104919; functional catalog ID=98"

/codon_start=1

/evidence=not_experimental

/product="putative RNaseP-associated protein"

/protein_id="AAD22641.1"

/db_xref="GI:4558548"

/translation="MRGVTPINPQKKIADSNPVKESDYEGERLTHLLDLIQRGIET SKLSNVNSLPKIKQIAIGINEVTVRLERMNPNTSDQQNPKQLQVILVADCK PRMLTKHIPNLAASRNVPVLYVRDNKRASRLGELVKLTALAIIGIKARGNDLNLK QILPRDS"

complement(14442. .19647)

/gene="T7B11.6"

/note="encodes putative SecA-type chloroplast protein transport factor; gene model last edited on 26 Mar 99"

/evidence=not_experimental

complement(join(14442. .14579,14663. .14902,14991. .15108,

CDS

13384. .14244

/gene="T7B11.5"

/note="encodes RNaseP-associated protein; identical to T15B16.20, GenBank accession number AF104919; gene model last edited on 3 Mar 99"

/evidence=not_experimental

join(13384. .13454,13566. .13613,13697. .13757,13842. .13940, 14020. .14244)

/gene="T7B11.5"

/note="identical to T15B16.20, GenBank accession number AF104919; functional catalog ID=98"

/codon_start=1

/evidence=not_experimental

/product="putative RNaseP-associated protein"

/protein_id="AAD22641.1"

/db_xref="GI:4558548"

/translation="MRGVTPINPQKKIADSNPVKESDYEGERLTHLLDLIQRGIET SKLSNVNSLPKIKQIAIGINEVTVRLERMNPNTSDQQNPKQLQVILVADCK PRMLTKHIPNLAASRNVPVLYVRDNKRASRLGELVKLTALAIIGIKARGNDLNLK QILPRDS"

complement(14442. .19647)

/gene="T7B11.6"

/note="encodes putative SecA-type chloroplast protein transport factor; gene model last edited on 26 Mar 99"

/evidence=not_experimental

complement(join(14442. .14579,14663. .14902,14991. .15108,

gene

8080. .10112

/gene="T7B11.3"

/note="encodes hypothetical protein; identical to T15B16.8, GenBank accession number AF104919; gene model last edited on 3 Mar 99"

/evidence=not_experimental

join(8080. .8568,9318. .9479,9557. .9778,9900. .10112)

/gene="T7B11.3"

/note="identical to T15B16.8, GenBank accession number AF104919; similar to T7B11.1; functional catalog ID=99"

/codon_start=1

/evidence=not_experimental

/product="hypothetical protein"

/protein_id="AAD22639.1"

/db_xref="GI:4558546"

/translation="MEQKQHILKQSTFSSPSSYSSISDRPISLLSRNGLLLLLLALV LLLGVLLPWPSPFLFPNRLSSSLSPSQSKWRDYTLAQAAKFVAKNGTVIVCAVSS PFLPFLNNWLISVSRQKHQKLVIAEDYITLYKVKWPGHVLIPPALDSKTAFSF GSQGFNFNTARRPQHLLQILELGYNVMYNDVMWLQDPFYLEGSHDAYFTDDMPQI KPLNHSDDLPHDRNGETIYICSCMIYLRPTNGAKLMLKWSSELQSAWSESIRFKAN DQPAFNALNKAHQVDLYLLSQVAFPTGGLYFNKNEAWVQETGKHVIVHNNYIIGYD RKMRFQDYGLWLVDHLESPGLKLE"

10826. .12762

/gene="T7B11.4"

/note="encodes hypothetical protein; identical to T15B16.7, GenBank accession number AF104919; gene model last edited on 3 Mar 99"

/evidence=not_experimental

join(10826. .10842,10929. .11371,11559. .11770,11843. .12130, 12260. .12433,12526. .12762)

/gene="T7B11.4"

/note="identical to T15B16.7, GenBank accession number AF104919; functional catalog ID=99"

/codon_start=1

/evidence=not_experimental

/product="hypothetical protein"

/protein_id="AAD22640.1"

/db_xref="GI:4558547"

/translation="MGTRHCVGLGYEMFNFWSSSVFVWSSSETLFKIVITRSLC ILGVHTLLTFKVLTFLLXIELQVFCVDCVVAETGIVVNIPTTKAQDGRPAGESGSK LRDEYILRLNGLLFDKAYRVDGKGKMMKXGPKGLGYGTAGADDYNDVGREGE ESAKLNAIQERTMRHIQIVDDHERLTKLLESEKKLEIKGNELAKPQVHNGTERMKL SEDLEORKEELHEKIIIRLERQIDQVQAIETEVEQLKGQLNVMMKMSDGDQAQVKEV DIIFKDLVEKBAELADLNKFNQTLILRERTNDELQEARKELVNMGELVRKPFVDM QQKYCQDVEDRAVEVLQWHEYINDPDWHPYKRKLENDREVEVIDDRDEKLRELK ADLGDGPNVATKALLEINEYNPSCRYITTELMNFEDKRAITLEEGVTCILDQWEKAK RKRGM"

13384. .14244

/gene="T7B11.5"

/note="encodes RNaseP-associated protein; identical to T15B16.20, GenBank accession number AF104919; gene model last edited on 3 Mar 99"

/evidence=not_experimental

join(13384. .13454,13566. .13613,13697. .13757,13842. .13940, 14020. .14244)

/gene="T7B11.5"

/note="identical to T15B16.20, GenBank accession number AF104919; functional catalog ID=98"

/codon_start=1

/evidence=not_experimental

/product="putative RNaseP-associated protein"

/protein_id="AAD22641.1"

/db_xref="GI:4558548"

/translation="MRGVTPINPQKKIADSNPVKESDYEGERLTHLLDLIQRGIET SKLSNVNSLPKIKQIAIGINEVTVRLERMNPNTSDQQNPKQLQVILVADCK PRMLTKHIPNLAASRNVPVLYVRDNKRASRLGELVKLTALAIIGIKARGNDLNLK QILPRDS"

complement(14442. .19647)

/gene="T7B11.6"

/note="encodes putative SecA-type chloroplast protein transport factor; gene model last edited on 26 Mar 99"

/evidence=not_experimental

complement(join(14442. .14579,14663. .14902,14991. .15108,

CDS

10826. .12762

/gene="T7B11.4"

/note="encodes hypothetical protein; identical to T15B16.7, GenBank accession number AF104919; gene model last edited on 3 Mar 99"

/evidence=not_experimental

join(10826. .10842,10929. .11371,11559. .11770,11843. .12130, 12260. .12433,12526. .12762)

/gene="T7B11.4"

/note="identical to T15B16.7, GenBank accession number AF104919; functional catalog ID=99"

/codon_start=1

/evidence=not_experimental

/product="hypothetical protein"

/protein_id="AAD22640.1"

/db_xref="GI:4558547"

/translation="MGTRHCVGLGYEMFNFWSSSVFVWSSSETLFKIVITRSLC ILGVHTLLTFKVLTFLLXIELQVFCVDCVVAETGIVVNIPTTKAQDGRPAGESGSK LRDEYILRLNGLLFDKAYRVDGKGKMMKXGPKGLGYGTAGADDYNDVGREGE ESAKLNAIQERTMRHIQIVDDHERLTKLLESEKKLEIKGNELAKPQVHNGTERMKL SEDLEORKEELHEKIIIRLERQIDQVQAIETEVEQLKGQLNVMMKMSDGDQAQVKEV DIIFKDLVEKBAELADLNKFNQTLILRERTNDELQEARKELVNMGELVRKPFVDM QQKYCQDVEDRAVEVLQWHEYINDPDWHPYKRKLENDREVEVIDDRDEKLRELK ADLGDGPNVATKALLEINEYNPSCRYITTELMNFEDKRAITLEEGVTCILDQWEKAK RKRGM"

13384. .14244

/gene="T7B11.5"

/note="encodes RNaseP-associated protein; identical to T15B16.20, GenBank accession number AF104919; gene model last edited on 3 Mar 99"

/evidence=not_experimental

join(13384. .13454,13566. .13613,13697. .13757,13842. .13940, 14020. .14244)

/gene="T7B11.5"

/note="identical to T15B16.20, GenBank accession number AF104919; functional catalog ID=98"

/codon_start=1

/evidence=not_experimental

/product="putative RNaseP-associated protein"

/protein_id="AAD22641.1"

/db_xref="GI:4558548"

/translation="MRGVTPINPQKKIADSNPVKESDYEGERLTHLLDLIQRGIET SKLSNVNSLPKIKQIAIGINEVTVRLERMNPNTSDQQNPKQLQVILVADCK PRMLTKHIPNLAASRNVPVLYVRDNKRASRLGELVKLTALAIIGIKARGNDLNLK QILPRDS"

complement(14442. .19647)

/gene="T7B11.6"

/note="encodes putative SecA-type chloroplast protein transport factor; gene model last edited on 26 Mar 99"

/evidence=not_experimental

complement(join(14442. .14579,14663. .14902,14991. .15108,

Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN,
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Yoshino,M. and Hayashizaki,Y.

TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from

Japonica rice

Science 301 (5631), 376-379 (2003)

JOURNAL MEDLINE

22752273

PUBMED

12869764

REFERENCE

AUTHORS

2 (bases 1 to 902)

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Oka,M., Ooka,H.,
Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Nomura,K.,
Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K.,
Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
Yoshimura,A.

Direct Submission

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT This clone is one of the 28K full-length cDNA clones from japonica
rice.

URL : <http://cdna01.dna.affrc.go.jp/cdna/>
NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K.,
Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y.,
Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N.,
Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,
Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,
Yasunishi,A. and Hayashizaki,Y.

Location/Qualifiers

1..902

/organism="Oryza sativa (japonica cultivar-group)"

FEATURES

source

/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J013149B08"

ORIGIN

Query Match 33.9%; Score 199.4; DB 8; Length 902;
Best Local Similarity 71.3%; Pred. No. 1.6e-39;
Matches 263; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

Qy 208 CCAGACTCGAAATTTTACAAAGTGGAGCAATTTGTCAGACCATGGAGATCCAGCAAGTT 267

Db 283 CCGGATCGGAGTTTCTACAAAGTGGAGCAATCTTGAGGCCATGGAGGTCCTTATGTG 342

Qy 268 TCATCGGCTTTACTGAAATCGGGATTCGAGGTGTTACTGTTTCTGATGTGAGGGTTT 327

Db 343 TCATCGGCTTTGCTGCAATGGGGATCAGAGCGGTGACGGTGTCCGACGTGCGGGTTTC 402

Qy 328 GGTGCAAGGAGGTTCTACCGAGAGACACGGTGGCTCTGAGTTCTCGGAAGACAAATTT 387

Db 403 GGCGCACAGGCGGGTCAACTGAGAGGCATGAAGGTCAGAATTTGCAGAAGATACATTT 462

Qy 388 GTTGCTAAAGTTAAAGTGGAAATCGTTGTTAAGAAAGACCAAGTGGATCTGTAAATCAAC 447

Db 463 ATTGATAAAGTTAAGATGGAATAGTGTGTCTCCAAGGATCAGGTTGAAGCTGTTGTGAC 522

Qy 448 ACAATAATTGAAGGAGCAAGGACAGGAGAGATGGTGATGGCAAGATTTTGTTCCT 507

Db 523 AGATAAATTGAAAAGGCAAGACAGGAGAAATTTGGTGATGGAAAAATATTTTGATACCC 582

Qy 508 GTGTCAGATGTCATAAGAGTTAGGACAGGTGAGCGTGGGAGAAAGCAGAGAAGATGACT 567

Db 583 GTGTCGGACGTGATCAGATACGCCACCGCGAACCAGGGGAGCGAGCGGAGAGGATGGCC 642

Qy 568 GGTGATATG 576

Db 643 GGAGGGCTG 651

RESULT 15

AK099152

LOCUS

DEFINITION

insert sequence.

ACCESSION

AK099152

VERSION

AK099152.1 GI:32984361

KEYWORDS

FLI_CDNA; CAP trapper.

SOURCE

Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartioideae; Oryzeae; Oryza.

REFERENCE

AUTHORS

1

The Rice Full-length cDNA Consortium, National Institute of

Agrobiological Sciences Rice Full-length cDNA Project Team;

Kikuchi,S., Satoh,K., Magata,T., Kawagashira,N., Doi,K.,

Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,

Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,

Ohtsuki,K., Shishiki,T., Foundation of Advancement of International

Science Genome Sequencing & Analysis Group; Otomo,Y., Murakami,K.,

Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,

Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,

Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,

Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,

Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN;

Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,

Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,

Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,

Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,

Yoshino,M. and Hayashizaki,Y.

TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from

japonica rice

Science 301 (5631), 376-379 (2003)

JOURNAL

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 19:07:55 ; Search time 1362.19 Seconds
(without alignments)
12890.278 Million cell updates/sec

Title: US-09-756-541-15
Perfect score: 588
Sequence: 1 ATGGCGCGTCAATGACGAA.....GTGATATGCTTTCACCGTCT 588

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estmu:*
 - 5: em_estov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_htc:*
 - 9: gb_est1:*
 - 10: gb_est2:*
 - 11: gb_htc:*
 - 12: gb_est3:*
 - 13: gb_est4:*
 - 14: gb_est5:*
 - 15: em_estfun:*
 - 16: em_estom:*
 - 17: em_gss_hum:*
 - 18: em_gss_inv:*
 - 19: em_gss_pln:*
 - 20: em_gss_vrt:*
 - 21: em_gss_fun:*
 - 22: em_gss_mam:*
 - 23: em_gss_mus:*
 - 24: em_gss_pro:*
 - 25: em_gss_rod:*
 - 26: em_gss_phg:*
 - 27: em_gss_vrl:*
 - 28: gb_gss1:*
 - 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	481	81.8	549	14 CB261070	CB261070 33-E9570-
2	474	80.6	566	9 AU236084	AU236084 AU236084
3	443.8	75.5	685	14 CD825321	CD825321 BN25.060G
4	442.2	75.2	669	14 CD822497	CD822497 BN25.045G

5	442.2	75.2	732	14	CD820903	CD820903 BN25.040D
6	440.6	74.9	745	14	CD819046	CD819046 BN20.047L
7	270.6	46.0	782	14	CB289081	CB289081 V-B-112F0
8	262.8	44.7	723	13	BQ798001	BQ798001 EST 6939
9	252.8	43.0	732	9	AJ558383	AJ558383 AJ558383
10	246.6	41.9	597	10	BE239550	BE239550 EST403599
11	246.6	41.9	714	12	BI311072	BI311072 EST531282
12	246.6	41.9	812	14	CA922829	CA922829 EST640547
13	245	41.7	613	9	AW035791	AW035791 EST281945
14	245	41.7	702	12	BI932123	BI932123 EST552012
15	245	41.7	711	12	BI932336	BI932336 EST552225
16	245	41.7	736	12	BI932913	BI932913 EST552802
17	245	41.7	741	10	AW738071	AW738071 EST339498
18	245	41.7	804	12	BI930336	BI930336 EST550225
19	245	41.7	810	12	BI930060	BI930060 EST549949
20	244.8	41.6	612	10	BF645522	BF645522 NF036F08E
21	244	41.5	648	9	AI773079	AI773079 EST254179
22	243.4	41.4	630	10	AW929534	AW929534 EST338322
23	243.4	41.4	685	12	BI929601	BI929601 EST549490
24	241.8	41.1	799	12	BM113473	BM113473 EST561009
25	240.8	41.0	677	10	BE823431	BE823431 GM700019B
26	240.2	40.9	756	12	BG595705	BG595705 EST494383
27	240.2	40.9	759	14	CF920467	CF920467 Gmrhrw3-
28	239.4	40.7	458	10	BE210153	BE210153 SO39d03.Y
29	231.8	39.4	560	14	CA514765	CA514765 KS09030G0
30	231.8	39.4	579	10	AW223643	AW223643 EST300454
31	224.8	38.2	675	14	CB346213	CB346213 CAB2SG000
32	223.2	38.0	719	12	BJ573686	BJ573686 BJ573686
33	222.6	37.9	866	13	C81819	C81819 C81819 Citr
34	222.4	37.8	540	14	CB349912	CB349912 CAB2SG000
35	221.8	37.7	551	14	CB349990	CB349990 CAB2SG000
36	221.8	37.7	598	14	CB350294	CB350294 CAB2SG000
37	221.6	37.7	657	14	CB349285	CB349285 CAB2SG000
38	221.6	37.7	701	14	CB349370	CB349370 CAB2SG000
39	219.2	37.3	540	10	AW624562	AW624562 EST322507
40	218.6	37.2	532	14	CB350222	CB350222 CAB2SG000
41	213	36.2	564	13	BQ118918	BQ118918 EST604494
42	213	36.2	729	14	CF475242	CF475242 RTWW2.14
43	212.8	36.2	796	14	CF395143	CF395143 RTDS2.9.G
44	211.8	36.0	673	14	CB085731	CB085731 hg15b09.9
45	210.8	35.9	713	13	BX248909	BX248909 BX248909

ALIGNMENTS

RESULT 1
CB261070
LOCUS
DEFINITION
33-E9570-012-004-A09-T7R MP1Z-ADIS-012 Arabidopsis thaliana cDNA
clone MP1Zp769A094Q 5-PRIME, mRNA sequence.
549 bp mRNA linear EST 06-NOV-2003

ACCESSION
CB261070.1 GI:32885843
VERSION
EST.
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1 (bases 1 to 549)

REFERENCE

AUTHORS
Schmid,K.J., Soerensen,T.R., Stracke,R., Torjek,O., Altmann,T.,
Mitchell-Olds,T. and Weisshaar,B.

TITLE
Large-scale identification and analysis of genome-wide
single-nucleotide polymorphisms for mapping in Arabidopsis thaliana

JOURNAL
Genome Res. 13 (6), 1250-1257 (2003)

MEDLINE
22683290

PUBMED
12799357

COMMENT

Contact: Weisshaar B

ADIS DNA core facility at MP1Z

Max-Planck-Institute for Plant Breeding Research

Carl-von-Linne Weg 10, 50829 Koeln, Germany

Fax: 00492215062851

Email: weisshaar@mpiz-koeln.mpg.de

Insert Length: 549 Std Error: 0.00	
Plate: 4 Row: A Column: 09	
Seq primer: T7R; CTAATACGACTCACTATAGGGA.	
FEATURES	Location/Qualifiers
	1..549
source	/organism="Arabidopsis thaliana"
	/mol_type="mRNA"
	/cultivar="Landsberg erecta (Ler)"
	/db_xref="GABI:589834"
	/db_xref="taxon:3702"
	/clone="MPIZp763A094Q"
	/tissue_type="whole plant"
	/dev_stage="adult plant, mixed stresses"
	/lab_host="E. coli XLI-Blue MRF"
	/clone_lib="MPIZ-ADIS-012"
	/note="Vector: pSPORT1; Site 1: Sali; Site 2: NotI; cDNA library from Arabidopsis thaliana, accession Landsberg erecta; six weeks old total plants grown under long-day conditions in soil, whole adult plants were treated for 24 hours with different stresses, (1) at 4M-0 C in the dark, (2), at 37 Grad C in the dark, (3) lying in the lab after removing from soil, (4) in the greenhouse after wounding leaves with a foreceps, (5) in the lab watering with a 150 mM NaCl solution, (6) at 26 Grad C in the light/UV; equal quantities of stressed plant material were pooled; library was made at the Max-Planck-Institute for Plant Breeding Research, Cologne, Germany; cloning sites Sali-NotI, primer sites and orientation:
	T7-Sali-CCACGCGTCCG-Sprime-cDNA-polyA-CC-NotI-SP6; Note: Sequencing granted in the context of the GABI Arabidopsis Verbund I: Genetic Diversity, 'Establishment of high-efficiency SNP-based mapping tools and development of methods for genome-wide mutation detection' PI: Bernd Weisshaar Sequence submission managed by RZPD/GABI-Primary database: http://gabi.rzpd.de. This clone is available from RZPD; contact RZPD (clone@rzpd.de) for further information."
ORIGIN	
Query Match 81.8%; Score 481; DB 14; Length 549;	
Best Local Similarity 97.7%; Pred. No. 1e-124;	
Matches 509; Conservative 0; Mismatches 10; Indels 2; Gaps 2;	
QY	1 ATGGCGGGCTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGA 60
Db	
QY	30 ATGGAGGGCTCATGACGAAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGA 89
Db	
QY	61 AAGAACATTGCTTCTCTGATTGCAATTCGATTGTTCTGATTTCAGACATTCGCGACCA 120
Db	
QY	90 AAG-ACATTGCTTCTCTGATTGCAATTCGATTGTTCTGATTTCAGACATTCGCGACCA 148
Db	
QY	121 TCTTGCTCGATTGGTCACAAAGTCACCGAGTAATAACAGTCGTCGTTTACCTGTCGTT 180
Db	
QY	149 TCITGCTCGATTGGTCACAAAGTCACCGAGTAATAACACTCGTGTGTTTACCTGTCGTT 208
Db	
QY	181 AGTGCCCAATATCTCTGATTATATTCAGACTC-GAAATTTTACAAGGTGGAAGCAAT 239
Db	
QY	209 AGAGCCCAAGCTCTTCTGATTATATTCAGACTCAGAAATTTTACAAGGTGGAAGCAAT 268
QY	240 TGTCAGACCATGGAGATCCAGCAAGTTTTCATCGGTTTACTGAAATTCGGATTTCGAGG 299
Db	
QY	269 TGTCAGACCATGGAGATCCAGCAAGTTTTCATCGGTTTACTGAAATTCGGATTTCGAGG 328
Db	
QY	300 TGTTACTGTTTCTGATGTGAGAGGGTTTGGTGACAGAGAGTTTCTACCGAGAGACACGG 359
Db	
QY	329 TGTTACTGTTTCTGATGTTCGAGGGTTTGGTGACAGAGAGTTTCTACTGAGAGACACGG 388
Db	
QY	360 TGGCTCTGATTCTCGGAAGACAAATTTGTTGCTAAGTTAAGATGGAATCGTTGTTAA 419
Db	
QY	389 TGGCTCTGATTCTCGGAAGACAAATTTGTTGCTAAGTTAAGATGGAATCGTTGTTAA 448
Db	
QY	420 GAAAGACCAAGTGGATCTGTAATCAACACAATAATTGAAGGAGCAAGGACAGGAGAT 479
Db	

Db	449 GAAAGACCAAGTGGATCTGTAATCAACACAATAATTGATGGAGCAAGGACAGGAGAT 508
QY	480 TGGTGATGGCAAGATTTTGTGTTTGCCTGCTCAGATGTCA 520
Db	
QY	509 TGGTGATGGCAAGATTTTGTGTTTGCCTGCTCAGATGTCA 549
Db	
RESULT 2	
AU236084	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
FEATURES	
source	
Location/Qualifiers	
1..566	
/organism="Arabidopsis thaliana"	
/mol_type="mRNA"	
/db_xref="taxon:3702"	
/clone="RAFL14-63-K15"	
/tissue_type="root"	
/lab_host="DH10B"	
/clone_lib="RAFL14"	
/note="Site_1: BamHI; Site_2: Sali"	
ORIGIN	
Query Match 80.6%; Score 474; DB 9; Length 566;	
Best Local Similarity 97.8%; Pred. No. 9.9e-123;	
Matches 500; Conservative 0; Mismatches 8; Indels 3; Gaps 2;	
QY	3 GCGGGCTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGAAA 62
Db	
QY	55 GCNGGCTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGAAA 114
Db	
QY	63 GAACATTGCTTCTCTGATTGCAATTCGATTGTTCTGGAATTCAGACATTCGACCATC 122
Db	
QY	115 GAACATTGCTTCTCTGATTGCAATTCGATTGTTCTGGAATTCAGACATTCGACCATC 174
Db	
QY	123 TTGCCTCGATTGGTCACAAAGTCACCGAGTAATAACAGTCGTCGTTTACCTGTCGTTAG 182
Db	
QY	175 TTGCCTCGATTGGTCACAAAGTCACCGAGTAATAACAGTCGTCGTTTACCTGTCGTTAG 234
Db	
QY	183 TGCCCAATATCTTCTGATTATATTCAGACTCGAAATTTTACAAGGTGGAAGCAATGT 242
Db	
QY	235 TGCCCAATATCTTCTGATTATATTCAGACTCGAAATTTTACAAGGTGGAAGCAATGT 294
Db	

QY 243 CAGACCATGGAGAAATCCAGCAAGTTTTCATCGGCTTTACTGAAATCCGGATTTCGAGGTGT 302
 Db 295 CAGACCATGGAGAAATCCAGCAAGTTTTCATCGGCTTTACTGAAATCCGGATTTCGAGGTGT 354
 QY 303 TACTGTTTCTGATGTGAGAGGGTTTGGTGCACAAGGAGTTTCTACCGAGAGACACGGTGG 362
 Db 355 TACTGTTTCTGATGTGAGAGGGTTTGGTGCACAAGGAGTTTCTACCGAGAGACACGGTGG 414
 QY 363 CTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAGAA 422
 Db 415 CTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAGAA 474
 QY 423 AGACCAAGT-GGAATCTGTAATCAACACAATAATTAAGGAGCAAGG--ACAGGAGAGAT 479
 Db 475 AGACCAAGTGGGAATCTGTAATCAACACAATAATTAAGGAGCCANGGACAGGGAGAGAT 534
 QY 480 TGGTGATGGCAAGATTTTGTGCTGTG 510
 Db 535 TGGTGATGGCAAGATTTTGTGCTGTG 565

RESULT 3
 CD825321
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Brassica napus (rape)
 Brassica napus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 685)
 Genopiante.
 Genopiante, a major partnership french program in plant genomics
 Unpublished (2003)
 Contact: Genopiante
 Genopiante
 93, rue Henri Rochefort 91025 EVRY CEDEX France
 Tel: 33 1 69 47 54 00
 Fax: 33 1 69 47 54 10
 This sequence has been generated in the framework of the french
 plant genomics programme 'Genopiante' (<http://www.genopiante.com>
 and <http://genopiante-info.infobiogen.fr>).
 Location/Qualifiers
 1..685
 /organism="Brassica napus"
 /mol_type="mRNA"
 /cultivar="Jet neuf"
 /db_xref="taxon:3708"
 /clone="BN25060G17"
 /tissue_type="seed"
 /clone_lib="BN25"

FEATURES
 source
 Query Match 75.5%; Score 443.8; DB 14; Length 685;
 Best Local Similarity 85.6%; Pred. No. 3.5e-114;
 Matches 506; Conservative 0; Mismatches 82; Indels 3; Gaps 1;
 QY 1 ATGGCGCGTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGA 60
 Db 9 ATGGCGCGTCAATGCGGAAACCCATCTCTGCAACCTCTCTCGGTTTCTACTCCGACCGA 68
 QY 61 AAG---AACATTGCTTTCTCTGATTGCAATTCGATTGTTGTTCTGGAATTCAGACATTCGCCGA 117
 Db 69 AAGAACAACATTTCTTCTCAGATTGCAATCTCGATATCTTCTGCGTTCCGACATTCGCCCA 128
 QY 118 CCATCTTGCTCGATTGTCACAAAGTCACCGAGTAATAACAGTCGTGTTTACCTGTC 177
 Db 129 CCGTCTCCCTCGATTGTCGCAAAACCCACCGACCAAGGCTCGCGTTTCCCATCTC 188

QY 178 GTTAGTGCCTCAATATCTTCTGATTATATTCAGACTCGAAATTTTACAAGGTGGAAGCA 237
 Db 189 GTTAGAGCCCAAGCTCTTATGATTATATTCAGACTCAAAGTTTACAAGGTGGAAGCA 248
 QY 238 ATTGTGAGACCATGGAGATCCAGCAAGTTTTCATCGGCTTTACTGAAATTCGGATTTCGA 297
 Db 249 ATTGTGAGGCCATGGAGATCCAGCATGTTTTCATCGGCTTTACTGAAATTTGGATTTCGA 308
 QY 298 GGTGTTACTGTTTCTGATGTGAGAGGGTTTGGTGCCACAAGGAGTTTCTACCGAGAGACAC 357
 Db 309 GGTGTGACCGTTTCTGATGTGAGAGGGTTTGGTGCCACAAGTGGATCCAAAGAGAGACAT 368
 QY 358 GGTGGCTCTGAGTTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTT 417
 Db 369 GGTGGGTCTGAGTTTCTCTGAAGACAAGTTGTTGCTAAAGTTAAGATGGAGATCGTTGTT 428
 QY 418 AAGAAAGACCAAGTGGATCTGTAATCAACACAATAATTAAGGAGCAAGGAGAGAG 477
 Db 429 AAGAAAGAACCAAGTGGAGTCTGTAATCAACCAATATCGATGGAGCAAGAACAGGAGAG 488
 QY 478 ATTGGTGATGGCAAGATTTTGTGCTGTGCTGTGATGATGTTTTCACCGTCT 537
 Db 489 ATCGGAGATGGCAAAATATTCTGTTTGGCTGTGTCAGATGTCATAAGGGTTTCGACAGGT 548
 QY 538 GAGCGTGGGAGAAAGCAGAGAGATGACTGGTGATGATGTTTTCACCGTCT 588
 Db 549 GAGCGTGGAGAACCAAGCAGAGAGATGACTGGTGACATGCTTTCGTCTCATCT 599

RESULT 4
 CD822497
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Brassica napus (rape)
 Brassica napus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 669)
 Genopiante.
 Genopiante, a major partnership french program in plant genomics
 Unpublished (2003)
 Contact: Genopiante
 Genopiante
 93, rue Henri Rochefort 91025 EVRY CEDEX France
 Tel: 33 1 69 47 54 00
 Fax: 33 1 69 47 54 10
 This sequence has been generated in the framework of the french
 plant genomics programme 'Genopiante' (<http://www.genopiante.com>
 and <http://genopiante-info.infobiogen.fr>).
 Location/Qualifiers
 1..669
 /organism="Brassica napus"
 /mol_type="mRNA"
 /cultivar="Jet neuf"
 /db_xref="taxon:3708"
 /clone="BN25045G20"
 /tissue_type="seed"
 /clone_lib="BN25"

FEATURES
 source
 Query Match 75.2%; Score 442.2; DB 14; Length 669;
 Best Local Similarity 85.4%; Pred. No. 9.8e-114;
 Matches 505; Conservative 0; Mismatches 83; Indels 3; Gaps 1;
 QY 1 ATGGCGCGTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGA 60
 Db 2 ATGGCGCGTCAATGCGGAAACCCATCTCTGCAACCTCTCTCGGTTTCTACTCCGACCGA 61

LOCUS	BQ798001	723 bp	mRNA	linear	EST 30-JUL-2002							
DEFINITION	EST 6939 Ripening Grape berries Lambda Zap II Library Vitis vinifera cDNA clone RT093C02 3', mRNA sequence.											
ACCESSION	BQ798001											
VERSION	BQ798001.1	GI:22012967										
KEYWORDS	EST.											
SOURCE	Vitis vinifera											
ORGANISM	Vitis vinifera											
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; Vitaceae; Vitis.											
AUTHORS	1 (bases 1 to 723) Abbal,P., Agasse,A., Ageorges,A., Atanassova,R., Barrieu,F., Couture,C., Dedaldechamp,F., Delrot,S., Glissant,D., Grimplet,J., Hamdi,S., Romieu,C. and Terrier,N.											
TITLE	Generation of Expressed Sequence Tag from Grape Berry (skin, pulp or seeds) at Various Developmental Stages											
JOURNAL	Unpublished (2002)											
COMMENT	Contact: Romieu C. Unite de Recherche des Produits de la Vigne Institut National de la Recherche Agronomique 2, place Viala, 34 060 Montpellier Cedex 01, France Tel: 00-33-(0)4-99-61-28-62 Fax: 00-33-(0)4-99-61-28-57 Email: romieu@ensam.inra.fr Seq primer: T7.											
FEATURES	Location/Qualifiers											
source	1..723											
	/organism="Vitis vinifera"											
	/mol_type="mRNA"											
	/cultivar="Shiraz"											
	/db_xref="taxon:29760"											
	/clone="RT093C02"											
	/dev_stage="ripening stage"											
	/clone_lib="Ripening Grape berries Lambda Zap II Library"											
	/note="Organ: Fruit; Vector: Lambda Zap II; Site 1: Eco RI; Site 2: XhoI; Oriented library, construction described in Generation of ESTs from grape Berry (skin, pulp or seeds) at various developmental stages by Terrier,N., Ageorges,A., Abbal,P., Romieu,C. in J. Plant Physiol. 158 (12): 1575-83 2001"											
ORIGIN												
Query Match	44.7%; Score 262.8; DB 13; Length 723;											
Best Local Similarity	76.8%; Pred. No. 4.4e-63;											
Matches	321; Conservative 0; Mismatches 97; Indels 0; Gaps 0;											
QY	154	AATAACAGTCGTGTTTACCTGTCGTAGTCCCAAAATATCTTCTGATTATATCCAGAC	213									
Db	675	AAGATGGATCTGTTCTCTCCATTGTTCAGAGCTCAGAGCTCTCCAGATTATACCCAGAC	616									
QY	214	TCGAAATTTTACAGGTGAAGCAATTGTCAGACCATGGAGAAATCCAGCAAGTTTCATCG	273									
Db	615	GCGCAGTTTACAAAGTTGAAGCGATCCTGAGGCCCTGCGGAATCCAGCAGTTTCTTCG	556									
QY	274	GCTTTACTGAAATCCGGATTTCGAGGTGTTACTGTTCTCGATGTGAGAGGGTTTGGTCA	333									
Db	555	GCTTTGCTGAAATGGGTATTTCGTGGTGTACTGTTTCTGATGTCCGGGGCTTTGGTGCT	496									
QY	334	CAAGGAGGTTCTACCGAGACACACGGTGGCTCTGAGTTTCTCGGAAGACAAATTTGTTCT	393									
Db	495	CAAGCGGTTTCGCCCTGAAAGACAGGCAGGCTCTGAATTTCTGAGGACAAATTTGTTCT	436									
QY	394	AAAGTTAAGATGGAATCGTTGTTTAAAGAACCAAGTCGAATCTGTAATCAACACAATA	453									
Db	435	AAAGTCAAATGGAGATTGTGGTGAGCAAGACCAGGTGGAAGCAGTAATTGACAGATC	376									
QY	454	ATTGAAGGAGCAAGGACAGGAGAGATTGGTGATGGCAAGATTTTGTGTCCTGTGTCA	513									
Db	375	AATGAGGTGGCAAGGACTGGAGAGATTGGTGATGGCAAGATCTTTTGGTGGCCCATATCG	316									
QY	514	GATGTCATAAGAGTTAGGACAGGTGAGCGTGGGGAGAAAGCAGAGAGATGACTGGTG	571									

LOCUS	AJ558383	732 bp	mRNA	linear	EST 12-JUN-2003							
DEFINITION	AJ558383 Antirrhinum majus whole plant Antirrhinum majus cDNA clone 018 1 07 f22, mRNA sequence.											
ACCESSION	AJ558383											
VERSION	AJ558383.1	GI:31660955										
KEYWORDS	EST.											
SOURCE	Antirrhinum majus (snapdragon)											
ORGANISM	Antirrhinum majus											
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Lamiales; Antirrhinaceae; Antirrhineae; Antirrhinum.											
AUTHORS	1 (bases 1 to 732) Zachgo,S., Stueber,K., Saedler,H., Sommer,H. and Schwarz-Sommer,Z.											
TITLE	Antirrhinum EST collection											
JOURNAL	Unpublished (2003)											
COMMENT	Contact: Schwarz-Sommer Z Molekulare Pflanzen-genetik MPI fuer Zuechtungs-forschung Carl-von-Linne Weg 10, D-50829, Germany.											
FEATURES	Location/Qualifiers											
source	1..732											
	/organism="Antirrhinum majus"											
	/mol_type="mRNA"											
	/db_xref="taxon:4151"											
	/clone="018_1_07_f22"											
	/tissue_type="whole plant"											
	/clone_lib="Antirrhinum majus whole plant"											
ORIGIN												
Query Match	43.0%; Score 252.8; DB 9; Length 732;											
Best Local Similarity	77.0%; Pred. No. 2.9e-60;											
Matches	308; Conservative 0; Mismatches 92; Indels 0; Gaps 0;											
QY	172	CCTGTCGTTAGTGCCAAATATCTTCTGATTATATTCAGACTCGAAATTTTACAAGTG	231									
Db	172	CCAATTGTTAGGGCTCAGATTCTCCAGATCCCTACCTGACTCCAAAGTCTCAAAAGTC	231									
QY	232	GAAGCAATTGTCAGACCATGGAGATCCAGCAAGTTTCATCGGCTTTTACTGAAATCGGG	291									
Db	232	GAAGCTATTCTAAGACCCCTGGAGGATCCACACAGTTTCCTCGGCCTTGTGAAATGGGA	291									
QY	292	ATTGAGGTGTTACTGTTTCTGATGTGAGAGGGTTTGGTGCAAGGAGGTTTCTACCGAG	351									
Db	292	ATTGTTGGTGTCACTGCTCTGATGTACGAGGGTTTGGTTCTCAGGGTGGTTTGACAGAA	351									
QY	352	AGACACGGTGGCTCTGAGTTCTCGGAGACAAATTTGTTGCTAAAGTTAAGATGGAATC	411									
Db	352	AGACAGTTTGGCTCTGAATTTCTCAGAGACAAATTTTGTGCAAAAGTTAAGATGGAGATT	411									
QY	412	GTTGTTAAGAAAGACCAAGTGGAAATCTGTAATCAACACAATTAATTGAAGGAGCAAGGACA	471									
Db	412	GTTGTGAGCAAAGACCAAGTTGATGCGGTAATAGAGAGATAATAGAGGAAGCAAGAACA	471									
QY	472	GGAGAGATTGGTGATGGCAAGATTTTGTGTTTTCCTGTTGTCAGATGTCATAAGAGTTAGG	531									
Db	472	GGAGAGATTGGTGATGGGAAGATTTTGTGCTTCCAGTTTCTGATGTAATAAGAGTTCGC	531									
QY	532	ACAGGTGAGCGTGGGGAGAAAGCAGAGAGATGACTGGTG	571									
Db	532	ACTGGTGAACGAGGGGAAAAGGCTGAGAGGATGGTGGGAG	571									
RESULT 10												
BE239550												
LOCUS	BE239550	597 bp	mRNA	linear	EST 12-JUL-2000							
DEFINITION	EST403599 MHRP- Medicago truncatula cDNA clone pMHRP-28F3, mRNA sequence.											

Db 172 TTCTTCCCAAAATCAGAGCTCRAAACCTTCCTGACTATGTTCTCTGAATCCAGTTTACA 231
Qy 227 AGGTGGAAGCAATTTGTGAGACCATCGAGATCCAGCAAGTTTTCATCGGCTTTACTGAAA 286
Db 232 AAGTTGAAGCCATTCTCAGGCCATGGAGAAATCCCTCAGGTTTCTTCGGGTTTGTGAAA 291
Qy 287 TCGGATTCGAGGTGTTACTGTTTCTGATGTGAGAGGGTTTGGTGCACAAGGAGTTCTA 346
Db 292 TGGGAATTCGTGGTGTCACTGTATCTGATGTCAAGGGTTTGGTGTCTCAGGTGGCTCA 351
Qy 347 CCGAGAGACACGGTGGCTCTGAGTTCTCGGAAGACAAAATTTGTTGCTAAAGTTAAGATGG 406
Db 352 AAGAGAGCAGGGAGGCTCCGAATTTTCTGAAGACAAATTTTGTGGCAAAAGTTAAATGG 411
Qy 407 AAATCGTTTGAAGAAAGACCAAGTGAATCTGTAATCAACACAAATAATGAAGGAGCAA 466
Db 412 AAATAGTGTGAGAAAGACCAAGTTGAGGCAGTGATACAAAATATGGAGACGGCAA 471
Qy 467 GGACAGAGAGATTTGGTGTGATGGCAAGATTTTGTGTTTGGCTGTGTCTCAGATGTCAAGAG 526
Db 472 GAACCTGGGAGATTTGGTGTGATGGCAAAATTTCTGTATCCCTGTATCTGATGTAATAGAA 531
Qy 527 TTAGGACAGGTGAGCGTGGGAGAAAGCAGAGAGATGACTGGTG 571
Db 532 TCCGACACAGTGAGCGTGGGAGACAGCGTGTGAGAGGATGGCTGGG 576

RESULT 12
CA922829/c
LOCUS
DEFINITION
sequence.
CA922829
CA922829.1 GI:27409759
EST:
SOURCE
ORGANISM
Medicago truncatula (barrel medic)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: VandenBosch K
Department of Plant Biology
University of Minnesota
220 BioSci Center, 1445 Gortner Ave, St. Paul, MN 55109, USA
Tel: 612 624 2755
Fax: 612 625 1738
Email: kvandenb@chb.umn.edu
Alias: Clone pMHRP-28F3
TIGR sequence name: MTUCF91TV
More information is available at: www.medicago.org
Seq primer: (gta Ata Cga Ctc Act Ata ggg C).

FEATURES
source
1. .812
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultiivar="Al7"
/db_xref="taxon:3880"
/clone="MTUS-58H7"
/tissue_type="mixed tissues"
/dev_stage="various stages"
/lab_host="XLOLR"
/clone_lib="MTUS"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging

extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XLOLR cells."

ORIGIN
Query Match 41.9%; Score 246.6; DB 14; Length 812;
Best Local Similarity 75.6%; Pred. No. 1.7e-58;
Matches 306; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
Qy 167 TTTTACCTGCTAGTGCCTCCAAATATCTCTGATATATATCCAGACTCGAAATTTTACA 226
Db 623 TTCTTCCCAAAATCAGAGCTCAAAACCTTCTGACTATGTTCTCTGAATCCAGTTTACA 564
Qy 227 AGGTGGAAGCAATTTGTGAGACCATGGAGATCCAGCAAGTTTTCATCGGCTTTACTGAAA 286
Db 563 AAGTTGAAGCCATTCTCAGGCCATGGAGAAATCCCTCAGGTTTCTTCGGGTTTGTGAAA 504
Qy 287 TCGGATTCGAGGTGTTACTGTTTCTGATGTGAGAGGGTTTGGTGCACAAGGAGTTCTA 346
Db 503 TGGGAATTCGTGGTGTCACTGTATCTGATGTCAAGGGTTTGGTGTCTCAGGTGGCTCAA 444
Qy 347 CCGAGAGACACGGTGGCTCTGAGTTCTCGGAGACAAATTTTGTGCTAAAGTTAAGATGG 406
Db 443 AAGAGAGGACGGAGGCTCCGAATTTTCTGAAGACAAATTTTGTGCTCAAAAGTTAAATGG 384
Qy 407 AAATCGTTTGAAGAAAGACCAAGTGAATCTGTAATCAACAAATAATGAAGGAGCAA 466
Db 383 AAATAGTGTGAGAAAGACCAAGTTGAGGCAGTGATACAAAATATGGAGACGGCAA 324
Qy 467 GGACAGGAGAGATTTGGTGTGATGGCAAGATTTTGTGTTTGGCTGTGTCTCAGATGTCAAGAG 526
Db 323 GAACTGGGGAGATTTGGTGTGATGGCAAAATTTTCTGTATCCCTGTATCTGATGTAATAGAA 264
Qy 527 TTAGGACAGGTGAGCGTGGGAGAAAGCAGAGAGATGACTGGTG 571
Db 263 TCCGACACAGGTGAGCGTGGGAGACAGCGTGTGAGAGGATGGCTGGGG 219

RESULT 13
AW035791
LOCUS
DEFINITION
613 bp mRNA linear EST 18-MAY-2001
EST281945 tomato callus, TAMU Lycopersicon esculentum cDNA clone
cLEC36D8, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Lycopersicon esculentum (tomato)
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 613)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
Holt,I.E., Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S.,
Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.
Generation of ESTs from tomato callus tissue
Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.
Location/Qualifiers
1. .613
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultiivar="TA496"
/db_xref="taxon:4081"
/clone="cLEC36D8"
/tissue_type="callus"
/dev_stage="25-40 days old"

/lab_host="XL1-Blue MRF"
/clone_lib="tomato callus, TAMU"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons
of seedlings 7-10 days post-germination were excised, cut
at both ends and placed on MS medium with no selection.
Mixed callus was harvested at 25 and 40 days and included
undifferentiated masses. Tomato Callus EST Library"

ORIGIN

Query Match 41.7%; Score 245; DB 9; Length 613;
Best Local Similarity 72.5%; Pred. No. 4.4e-58;
Matches 317; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 135 GGTCAAAAGTCACCGAGTAATAACAGTCGTGTTTACCTGTCGTAGTGCCCAATATC 194
|||||
Db 131 GGTCACTTCAAACGCTGTCAAATGCTCTCTTTCCCGATTATCAGAGCCCAAACTC 190
|||||

QY 195 TTCTGATTATATCCAGACTCGAAATTTTACAAGGTGGAGCAATGTCAGACCATGGAG 254
|||||
Db 191 TCCAGATTTCGTCCCTGATGCGAAGTTTACAAAGTTGAAGCAATCTAAGACCTTGGAG 250
|||||

QY 255 AATCCAGCAAGTTTCATCGGCTTTACTGAAATCGGGAATCCAGGTGTTACTGTTCTGA 314
|||||
Db 251 AATTCAACAGGTTTCTTCGGCACTACTGAAATGGCAATTCGTGTTCACTGTTTCGGA 310
|||||

QY 315 TGTGAGAGGTTTGTGTCACACAGGAGTTCTACCGAGAGACACGTTGGCTCTGAGTTCTC 374
|||||
Db 311 TGTTCGTGTTTGGCGCCCAAGTGGCTTGACTGAGAGGCAAGCTGGCTCTGAATCTC 370
|||||

QY 375 GGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAGAAAGACCAAGTGA 434
|||||
Db 371 TGAAGACACGTTTGTGTCAAAAGTTAAATGGAATTTGTTGTCAGCAAGACCAAGTTGA 430
|||||

QY 435 ATCTGTAATCAACACAATAATTGAAGAGCAAGGACAGGAGAGATTGGTGATGGCAAGAT 494
|||||
Db 431 AGGAGTCATTGCCATGATAATTGAAGAGGCAAGAACTGGTGAATAGGTGATGGAAGAT 490
|||||

QY 495 TTTTGTGTTGCTGTGTCAGATGTCATAAGAGTTAGGACAGGTGAGCGTGGGAGAAAGC 554
|||||
Db 491 ATTCTTGACTCCCATCTCCGATGTTATAAGAGTTGCACTGGTGAACGGGAGAAAGGC 550
|||||

QY 555 AGAGAAGATGACTGGTG 571
|||||
Db 551 AGAGAGGATGATGGGAG 567
|||||

RESULT 14
BI932123
LOCUS
DEFINITION
EST552012 tomato flower, 8 mm to preanthesis buds Lycopersicon
esculentum cDNA clone cTCC22G23 5' end, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Lycopersicon esculentum (tomato)
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 702)

REFERENCE
AUTHORS
van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
Generation of ESTs from tomato flower tissue, buds 8 mm -
preanthesis
Unpublished (2001)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
This clone is available through the Clemson University Genomics

Institute
Seq primer: T3.
Location/Qualifiers
1..702
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTCC22G23"
/tissue_type="flower"
/dev_stage="buds 8mm to preanthesis"
/clone_lib="tomato flower, 8 mm to preanthesis buds"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research; Flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

FEATURES
source

ORIGIN

Query Match 41.7%; Score 245; DB 12; Length 702;
Best Local Similarity 72.5%; Pred. No. 4.6e-58;
Matches 317; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 135 GGTCAAAAGTCACCGAGTAATAACAGTCGTGTTTACCTGTCGTAGTGCCCAATATC 194
|||||
Db 172 GGTCACTTCAAACGCTGTCAAATGCTCTCTTTCCCGATTATCAGAGCCCAAACTC 231
|||||

QY 195 TTCTGATTATATCCAGACTCGAAATTTTACAAGGTGGAGCAATGTCAGACCATGGAG 254
|||||
Db 232 TCCAGATTTCGTCCCTGATGCGAAGTTTACAAAGTTGAAGCAATCTAAGACCTTGGAG 291
|||||

QY 255 AATCCAGCAAGTTTCATCGGCTTTACTGAAATCGGGAATTCGAGGTGTTACTGTTCTGA 314
|||||
Db 292 AATTCAACAGGTTTCTTCGGCACTACTGAAATGGCAATTCGTTGTTCACTGTTTCGGA 351
|||||

QY 315 TGTGAGAGGTTTGGTGCACAAGGAGTTCTACCGAGAGACACGTTGGTCTCTGAGTTCTC 374
|||||
Db 352 TGTTCGTGTTTGGCGCCCAAGTGGCTTGACTGAGAGGCAAGCTGGCTCTGAATCTC 411
|||||

QY 375 GGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAGAAAGACCAAGTGA 434
|||||
Db 412 TGAAGACACGTTTGTGTCAAAAGTTAAATGGAATTTGTTGTCAGCAAGACCAAGTTGA 471
|||||

QY 435 ATCTGTAATCAACACAATAATTGAAGAGCAAGGACAGGAGAGATTGGTGATGGCAAGAT 494
|||||
Db 472 AGGAGTCATTGCCATGATAATTGAAGAGGCAAGAACTGGTGAATAGGTGATGGAAGAT 531
|||||

QY 495 TTTTGTGTTGCTGTGTCAGATGTCATAAGAGTTAGGACAGGTGAGCGTGGGAGAAAGC 554
|||||
Db 532 ATTCTTGACTCCCATCTCCGATGTTATAAGAGTTGCACTGGTGAACGGGAGAAAGGC 591
|||||

QY 555 AGAGAAGATGACTGGTG 571
|||||
Db 592 AGAGAGGATGATGGGAG 608
|||||

RESULT 15
BI932336
LOCUS
DEFINITION
EST552225 tomato flower, 8 mm to preanthesis buds Lycopersicon
esculentum cDNA clone cTCC22H10 5' end, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Lycopersicon esculentum (tomato)
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 711)

REFERENCE
AUTHORS
van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 13:28:59 ; Search time 205.069 Seconds
(without alignments)
12180.972 Million cell updates/sec

Title: US-09-756-541-15
Perfect score: 588
Sequence: 1 ATGGCGGCGTCAATGACGAA.....GTGATATGCTTTCACCGTCT 588

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB ID	Description
		Match	Length		
1	588	100.0	588	AAF58583	Aaf58583 Arabidops
2	588	100.0	591	ABZ14328	Abz14328 Arabidops
3	588	100.0	817	AAF58581	Aaf58581 Arabidops
4	285	48.5	594	AAF58584	Aaf58584 Ricinus c
5	285	48.5	897	AAF58582	Aaf58582 Ricinus c
6	92.8	15.8	96109	AAF28548	Aaf28548 Genomic f
7	85.8	14.6	375	ADA30024	Ada30024 DNA encod
8	84.6	14.4	110000	2 AAV21209_01	Continuation (2 of
9	84.4	14.4	110000	2 AAV21209_14	Continuation (15 o
10	81.4	13.8	339	7 ACF72027	Acf72027 Photorhab
11	81.4	13.8	110000	7 ACF67367_52	Continuation (53 o
12	81.4	13.8	110000	7 ACF65387_3	Continuation (4 of
13	79.8	13.6	110000	2 AAT42063_03	Continuation (4 of
14	77	13.1	363	7 ABZ40022	Abz40022 N. gonorr
15	73	12.4	16526	3 AAA81472	Aaa81472 N. mening
16	73	12.4	110000	3 AAA81489_6	Continuation (7 of
17	73	12.4	172325	3 AAF21613	Aaf21613 Neisseria
18	73	12.4	349980	3 AAF21612	Aaf21612 Neisseria
19	68.8	11.7	336	5 AAH68445	Aah68445 C glutami
20	68.8	11.7	417	5 AAH67222	Aah67222 C glutami
21	68.8	11.7	349980	5 AAH68531	Aah68531 C glutami
22	67.2	11.4	1164	5 AAS88441	Aas88441 DNA encod
23	67.2	11.4	11095	4 AAS46256	Aas46256 DNA encod

24	64.6	11.0	663	7 ACF65588	Acf65588 Photorhab
25	64.6	11.0	110000	7 ACF67367_46	Continuation (47 o
26	64.6	11.0	110000	7 ACF65388_01	Continuation (2 of
27	61	10.4	339	7 ACF71482	Acf71482 Photorhab
28	60.4	10.3	110000	6 ABQ69245_15	Continuation (16 o
29	60.4	10.3	110000	6 ABQ67195_1	Continuation (2 of
30	60.4	10.3	110000	6 ABQ67195_2	Continuation (3 of
31	59.6	10.1	290	4 AAF71296	Aaf71296 Corynebac
32	57.6	9.8	110000	6 ABA90521_16	Continuation (17 o
33	52.4	8.9	2751	6 ABQ70750	Abq70750 Listeria
34	50.8	8.6	110000	6 ABA03041_15	Continuation (16 o
35	49.6	8.4	110000	4 AAI99682_32	Continuation (33 o
36	49.6	8.4	110000	4 AAI99683_32	Continuation (33 o
37	47	8.0	342	6 ABN18985	Abn18985 Human ORF
38	44.4	7.6	174	7 ABZ39932	Abz39932 N. gonorr
39	43.4	7.4	3858	5 AAS89065	Aas89065 DNA encod
40	43.4	7.4	3858	5 AAS90406	Aas90406 DNA encod
41	43.4	7.4	3858	5 AAS85755	Aas85755 DNA encod
42	43.4	7.4	3858	5 AAS92604	Aas92604 DNA encod
43	43	7.3	2000	7 ADA71938	Ada71938 Rice gene
44	42.8	7.3	1970	3 AAZ99179	Aaz99179 T. hydrot
45	42.2	7.2	674	7 ACF67052	Acf67052 Photorhab

ALIGNMENTS

RESULT 1
AAF58583
ID AAF58583 standard; cDNA; 588 BP.
XX
AC AAF58583;
XX
DT 23-APR-2001 (first entry)
XX
DE Arabidopsis thaliana P-PII cDNA fragment.
XX
KW Arabidopsis thaliana; PII; plant nitrogen regulatory gene; P-PII;
KW nitrogen assimilation; transgenic plant; herbicide screening; ss.
XX
OS Arabidopsis thaliana.
XX
PN US6177275-B1.
XX
PD 23-JAN-2001.
XX
PF 23-JUL-1997; 97US-00899330.
XX
PR 24-JUL-1996; 96US-0022328P.
XX
PA (U'NY) UNIV NEW YORK STATE.
XX
PI Coruzzi GM, Lam H, Hsieh M;
XX
DR WPI; 2001-158572/16.
DR P-PSDB; AAB69495.
XX
PT Novel P-PII genes capable of regulating plant nitrogen assimilation,
PT useful for transgenic plant production, and as probes for isolating
PT additional genomic clones having P-PII gene promoters.
XX
PS Claim 2; Col 37-38; 35pp; English.
XX
CC The present sequence encodes a nitrogen regulatory PII protein. Novel
CC plant PII (also called P-PII) nucleotide sequences have been isolated.
CC They are useful for regulating nitrogen assimilation in plants, and in
CC transgenic plant production. They are also used to engineer organisms
CC that overexpress wild-type or mutant P-PII regulatory proteins. P-PII
CC proteins are useful for in vitro screening of herbicides. P-PII
CC nucleotides may be used as probes for isolating additional genomic clones
CC with the promoters of P-PII genes. P-PII promoters are light- and/or
CC sucrose-inducible, and are suitable for genetic engineering of plants

SQ Sequence 588 BP; 164 A; 105 C; 147 G; 172 T; 0 U; 0 Other;

Query Match 100.0%; Score 588; DB 4; Length 588;
Best Local Similarity 100.0%; Pred. No. 2.3e-168;
Matches 588; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGCGTCAATGACGAAACCCATCTCAATAACTTCTCTCGTTTCTATTCTGATCGA 60
DB 1 ATGGCGGCGTCAATGACGAAACCCATCTCAATAACTTCTCTCGTTTCTATTCTGATCGA 60

QY 61 AAGAACATTGCTTCTCTGATTGCAATTCGATTGTTCTCGATTTCAGACATTCGCGACCA 120
DB 61 AAGAACATTGCTTCTCTGATTGCAATTCGATTGTTCTCGATTTCAGACATTCGCGACCA 120

QY 121 TCTTGCCCTCGATTGGTCACAAAGTCACCGAGTAATAACAGTCGCTGTTTACCTGTCGTT 180
DB 121 TCTTGCCCTCGATTGGTCACAAAGTCACCGAGTAATAACAGTCGCTGTTTACCTGTCGTT 180

QY 181 AGTGCCCAATATCTTCTGATTATATTCAGACTCGAATTTTACAGGTGGAAGCAATT 240
DB 181 AGTGCCCAATATCTTCTGATTATATTCAGACTCGAATTTTACAGGTGGAAGCAATT 240

QY 241 GTCAGACCATGGAGAAATCCAGCAAGTTTCATCGGCTTTACTGAAATCGGGATTGAGGT 300
DB 241 GTCAGACCATGGAGAAATCCAGCAAGTTTCATCGGCTTTACTGAAATCGGGATTGAGGT 300

QY 301 GTTACTGTTTCTGATGTGAGAGGGTTTGGTGCAAGAGGTTCACCGAGAGACACGGT 360
DB 301 GTTACTGTTTCTGATGTGAGAGGGTTTGGTGCAAGAGGTTCACCGAGAGACACGGT 360

QY 361 GGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAG 420
DB 361 GGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAG 420

QY 421 AAGACCAAGTGAATCTGTAATCAACACAATAATTGAAGGAGCAAGGAGAGAGATT 480
DB 421 AAGACCAAGTGAATCTGTAATCAACACAATAATTGAAGGAGCAAGGAGAGAGATT 480

QY 481 GGTGATGGCAAGATTTTGTGTTGCTGTCAGATGTCATAAGAGTTAGGACAGGTGAG 540
DB 481 GGTGATGGCAAGATTTTGTGTTGCTGTCAGATGTCATAAGAGTTAGGACAGGTGAG 540

RESULT 2
ABZ14328
ID ABZ14328 standard; DNA; 591 BP.
XX AC ABZ14328;
XX DT 21-JAN-2003 (first entry);
XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2133.
XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX OS Arabidopsis thaliana.
XX PN WO200216655-A2.
XX PD 28-FEB-2002.
XX PF 24-AUG-2001; 2001WO-US026685.
XX PR 24-AUG-2000; 2000US-0227866P.
XX PR 26-JAN-2001; 2001US-0264647P.
XX PR 22-JUN-2001; 2001US-0300111P.
XX PA (SCRI) SCRIPPS RES INST.
XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX PI Harper JF, Kreps J, Wang X, Zhu T;
XX DR WPI; 2002-304127/34.
XX PT Identifying a stress condition to which a plant cell has been exposed and
XX PT producing plants with increased tolerance to these abiotic stresses.
XX PS Claim 144; SEQ ID NO 2133; 577pp + Sequence Listing; English.
XX CC The invention relates to identifying a stress condition to which a plant
XX CC cell has been exposed, comprising: (a) contacting nucleic acid
XX CC representative of expressed polynucleotides in the plant cell with an
XX CC array or probes representative of the plant cell genome; and (b)
XX CC detecting a profile of expressed polynucleotides in the plant cell
XX CC characteristic of a stress response. The method is useful in the
XX CC production of transgenic plants, cells and seeds and in producing plants
XX CC with increased tolerance to abiotic stress. The present sequence is that
XX CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
XX CC in methods of the invention. Note: The sequence data for this patent is
XX CC not represented in the printed specification but is based on sequence
XX CC information supplied to Derwent by the European Patent Office

SQ Sequence 591 BP; 165 A; 105 C; 148 G; 173 T; 0 U; 0 Other;

Query Match 100.0%; Score 588; DB 6; Length 591;
Best Local Similarity 100.0%; Pred. No. 2.3e-168;
Matches 588; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGCGTCAATGACGAAACCCATCTCAATAACTTCTCTCGTTTCTATTCTGATCGA 60
DB 1 ATGGCGGCGTCAATGACGAAACCCATCTCAATAACTTCTCTCGTTTCTATTCTGATCGA 60

QY 61 AAGAACATTGCTTCTCTGATTGCAATTCGATTGTTCTGATTTCAGACATTCGCGACCA 120
DB 61 AAGAACATTGCTTCTCTGATTGCAATTCGATTGTTCTGATTTCAGACATTCGCGACCA 120

QY 121 TCTTGCCCTCGATTGGTCACAAAGTCACCGAGTAATAACAGTCGCTGTTTACCTGTCGTT 180
DB 121 TCTTGCCCTCGATTGGTCACAAAGTCACCGAGTAATAACAGTCGCTGTTTACCTGTCGTT 180

QY 181 AGTGCCCAATATCTTCTGATTATATTCAGACTCGAATTTTACAGGTGGAAGCAATT 240
DB 181 AGTGCCCAATATCTTCTGATTATATTCAGACTCGAATTTTACAGGTGGAAGCAATT 240

QY 241 GTCAGACCATGGAGAAATCCAGCAAGTTTCATCGGCTTTACTGAAATCGGGATTGAGGT 300
DB 241 GTCAGACCATGGAGAAATCCAGCAAGTTTCATCGGCTTTACTGAAATCGGGATTGAGGT 300

QY 301 GTTACTGTTTCTGATGTGAGAGGGTTTGGTGCAAGAGGTTCACCGAGAGACACGGT 360
DB 301 GTTACTGTTTCTGATGTGAGAGGGTTTGGTGCAAGAGGTTCACCGAGAGACACGGT 360

QY 361 GGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAG 420
DB 361 GGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAG 420

QY 421 AAGACCAAGTGAATCTGTAATCAACACAATAATTGAAGGAGCAAGGAGAGAGATT 480
DB 421 AAGACCAAGTGAATCTGTAATCAACACAATAATTGAAGGAGCAAGGAGAGAGATT 480

QY 481 GGTGATGGCAAGATTTTGTGTTGCTGTCAGATGTCATAAGAGTTAGGACAGGTGAG 540
DB 481 GGTGATGGCAAGATTTTGTGTTGCTGTCAGATGTCATAAGAGTTAGGACAGGTGAG 540

QY 541 CGTGGGAGAAAGCAGAGAGAGATGACTGGTGATATGCTTTCACCGTCT 588
DB 541 CGTGGGAGAAAGCAGAGAGAGATGACTGGTGATATGCTTTCACCGTCT 588

RESULT 3
AAF58581
ID AAF58581 standard; cDNA; 817 BP.

XX AAF58581;
AC
XX 23-APR-2001 (first entry)
DT
XX Arabidopsis thaliana P-PII cDNA.
DE
XX Arabidopsis thaliana; PII; plant nitrogen regulatory gene; P-PII;
XW nitrogen assimilation; transgenic plant; herbicide screening; ss.
XW
XX Arabidopsis thaliana.
OS
XX US6177275-B1.
PN
XX 23-JAN-2001.
PD
XX 23-JUL-1997; 97US-00899330.
PF
XX 24-JUL-1996; 96US-0022328P.
PR
XX (UYNV) UNIV NEW YORK STATE.
PA
XX Coruzzi GM, Lam H, Hsieh M;
PI
XX WPI; 2001-158572/16.
DR
DR P-PSDB; AAB69495.
XX
XX Novel P-PII genes capable of regulating plant nitrogen assimilation,
PT useful for transgenic plant production, and as probes for isolating
PT additional genomic clones having P-PII gene promoters.
XX
XX Claim 2; Fig 12; 35pp; English.
XX
XX The present sequence encodes a nitrogen regulatory PII protein. Novel
CC plant PII (also called P-PII) nucleotide sequences have been isolated.
CC They are useful for regulating nitrogen assimilation in plants, and in
CC transgenic plant production. They are also used to engineer organisms
CC that overexpress wild-type or mutant P-PII regulatory proteins. P-PII
CC proteins are useful for in vitro screening of herbicides. P-PII
CC nucleotides may be used as probes for isolating additional genomic clones
CC with the promoters of P-PII genes. P-PII promoters are light- and/or
CC sucrose-inducible, and are suitable for genetic engineering of plants
XX
SQ Sequence 817 BP; 242 A; 137 C; 188 G; 250 T; 0 U; 0 Other;

Query Match 100.0%; Score 588; DB 4; Length 817;
Best Local Similarity 100.0%; Pred. No. 2.6e-168;
Matches 588; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGCGTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTCTATTCTGATCGA 60
Db
33 ATGGCGGCGTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTCTATTCTGATCGA 92

QY 61 AAGAACATTGCTTTCTGATTGCAATTTCGATTGTTCTCGGATTCAACATCCCGACCA 120
Db
93 AAGAACATTGCTTTCTGATTGCAATTTCGATTGTTCTCGGATTCAACATCCCGACCA 152

QY 121 TCTTGCCCTCGATTGGTCACAAAGTCACCGAGTAATAACAGTCGTGTTTACCTGTCGTT 180
Db
153 TCTTGCCCTCGATTGGTCACAAAGTCACCGAGTAATAACAGTCGTGTTTACCTGTCGTT 212

QY 181 AGTGCCCAATATCTTCTGATTATATTCAGACTCGAAATTTTACAAGGTGGAAGCAATT 240
Db
213 AGTGCCCAATATCTTCTGATTATATTCAGACTCGAAATTTTACAAGGTGGAAGCAATT 272

QY 241 GTCAGACCATGGAGATCCAGCAAGTTTCATCGGCTTTACTGAAATCGGATTTCGAGGT 300
Db
273 GTCAGACCATGGAGATCCAGCAAGTTTCATCGGCTTTACTGAAATCGGATTTCGAGGT 332

QY 301 GTTACTGTTTCTGATGTGAGAGGGTTTGGTGCAACAAGGAGTTCTACCGAGAGACACCGT 360
Db
333 GTTACTGTTTCTGATGTGAGAGGGTTTGGTGCAACAAGGAGTTCTACCGAGAGACACCGT 392

QY 361 GGCTCTGAGTTCTCGAAGACAAATTTGTTGCTAAAGTTAAGATGGAAATCGTTGTTAAG 420
Db
393 GGCTCTGAGTTCTCGAAGACAAATTTGTTGCTAAAGTTAAGATGGAAATCGTTGTTAAG 452

QY 421 AAAGACCAAGTGGAAATCTGTAATCAACACATAATTAAGGAGCAAGGACAGGAGATT 480
Db
453 AAAGACCAAGTGGAAATCTGTAATCAACACATAATTAAGGAGCAAGGACAGGAGATT 512

QY 481 GGTGATGGCAAGATTTTGTGTTTGCCTGTGTCAGATGTCATAAGATTAGGACAGGTGAG 540
Db
513 GGTGATGGCAAGATTTTGTGTTTGCCTGTGTCAGATGTCATAAGATTAGGACAGGTGAG 572

QY 541 CGTGGGAGAAAGCAGAGAAGATGACTGGTGATATGCTTTTCAACCGTCT 588
Db
573 CGTGGGAGAAAGCAGAGAAGATGACTGGTGATATGCTTTTCAACCGTCT 620

RESULT 4
AAF58584
ID AAF58584 standard; cDNA; 594 BP.
XX
AC AAF58584;
XX
DT 23-APR-2001 (first entry)
XX
DE Ricinus communis P-PII cDNA fragment.
XX
XW Castor bean; PII; plant nitrogen regulatory gene; P-PII;
KW nitrogen assimilation; transgenic plant; herbicide screening; ss.
XX
OS Ricinus communis.
XX
PN US6177275-B1.
XX
PD 23-JAN-2001.
XX
PF 23-JUL-1997; 97US-00899330.
XX
PR 24-JUL-1996; 96US-0022328P.
XX
PA (UYNV) UNIV NEW YORK STATE.
XX
PI Coruzzi GM, Lam H, Hsieh M;
XX
DR WPI; 2001-158572/16.
DR
XX P-PSDB; AAB69496.

PT Novel P-PII genes capable of regulating plant nitrogen assimilation,
PT useful for transgenic plant production, and as probes for isolating
PT additional genomic clones having P-PII gene promoters.
XX
PS Claim 2; Col 37-38; 35pp; English.
XX
CC The present sequence encodes a nitrogen regulatory PII protein. Novel
CC plant PII (also called P-PII) nucleotide sequences have been isolated.
CC They are useful for regulating nitrogen assimilation in plants, and in
CC transgenic plant production. They are also used to engineer organisms
CC that overexpress wild-type or mutant P-PII regulatory proteins. P-PII
CC proteins are useful for in vitro screening of herbicides. P-PII
CC nucleotides may be used as probes for isolating additional genomic clones
CC with the promoters of P-PII genes. P-PII promoters are light- and/or
CC sucrose-inducible, and are suitable for genetic engineering of plants
XX
SQ Sequence 594 BP; 149 A; 117 C; 157 G; 171 T; 0 U; 0 Other;

Query Match 48.5%; Score 285; DB 4; Length 594;
Best Local Similarity 70.3%; Pred. No. 3.4e-76;
Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

QY 7 GCGTCAATGACGAAACCCATCTCAATAACTTCTCTCGSTTTCTATTCTGATCGAAGAAC 66
Db
7 GAGGCTACTGCGAAACTGGGCTTGCTCACTCTCTTCTATTCTAATAACATCAAGAAAGAA 66

XX 28-DEC-2000.
PD 16-JUN-2000; 2000WC-US016649.
XX 22
PR 18-JUN-1999; 99US-0140121P.
XX (INCY-) INCYTE GENOMICS INC.
PA Lagace RE, Patterson C, Berg KL;
XX WPI; 2001-041427/05.
DR Genomic library for identifying diagnostic and therapeutic compositions,
XX and for identifying virulence factors, regulatory elements and drug
PT targets, comprises Moraxella catarrhalis nucleic acids.
PT
XX Claim 1; Page 345-368; 545pp; English.
PS The present invention relates to a Moraxella catarrhalis genomic library
XX comprising of a combination of 41 nucleic acid molecules (see AAF28514-
CC AAF28554). The library has a number of uses described in the
CC specification e.g. is useful for identifying diagnostic and therapeutic
CC compositions. M. catarrhalis (Branhamella catarrhalis) is a large
CC aerobic, gram-negative diplococcus, normally found among the bacterial
CC flora of human upper airways. M. catarrhalis is known to cause acute,
CC localised infections such as otitis media, sinusitis and bronchopulmonary
CC infection and life-threatening, systemic diseases including endocarditis
CC and meningitis
XX
SQ Sequence 96109 BP; 28783 A; 18910 C; 20341 G; 28075 T; 0 U; 0 Other;
Query Match 15.8%; Score 92.8; DB 4; Length 96109;
Best Local Similarity 57.8%; Pred. No. 8.8e-17;
Matches 185; Conservative 0; Mismatches 132; Indels 3; Gaps 1;
QY 235 GCAATTGTCTAGACCATGGAGATCCAGCAAGTTTTCATCGGCTTTACTGAAATCGGGATT 294
Db 12953 GCAATTATCAAGCCGTTTAAACTCGATGATGTCGAGAGCACTCTCAGAAATTGGCGTC 13012
QY 295 CGAGTGTCTACTGTTTCTGATGTGAGAGGTTTGGTGCAAGAGGTTCTACCGAGAGA 354
Db 13013 AATGTATCACCGTCACTCAAGTCAAGGCTTGGTCGCCAAAAGGTCATACCGAGATG 13072
QY 355 CACGTGGCTCTGAGTTCTCGGAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTT 414
Db 13073 TATCGTGGGCGGAATATGTGGTTGAT---TTTACCACAAATTAATAATTGAGATAGCA 13129
QY 415 GTTAAGAAAGACCAAGTGGATCTGTATCAACACACAATAATTGAAGGACAGGACAGGA 474
Db 13130 TGTCGTGATGAGATGGTTGATTCAATTATGAGTCAATCATTAAGTTGCAAAATACAGGT 13189
QY 475 GAGATTGGTGATGGCAAGATTTTGGTTTTCGCTGTGTCAGATGTCATAAGAGTTAGGACA 534
Db 13190 AAAATTGGTGATGGTAAGATTTTGGTTAGTCCGCTTGAGCGGTGTCATTCCGATTGAACT 13249
QY 535 GGTGAGCGTGGGAGAAAGC 554
Db 13250 GGCGAATTGATGAAAGTGC 13269
RESULT 7
ID ADA30024
XX ADA30024;
AC ADA30024;
XX
DT 20-NOV-2003 (first entry)
XX
DE DNA encoding Acinetobacter baumannii protein #1311.
XX ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;
KW vaccine; plant biocontrol agent.

XX Acinetobacter baumannii.
OS
XX US6562958-B1.
PN
XX 13-MAY-2003.
PD
XX 04-JUN-1999; 99US-00328352.
PF
XX 09-JUN-1998; 98US-0088701P.
PR
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX Breton G, Bush D;
PI WPI; 2003-576092/54.
XX P-PSDB; ADA34150.
DR
XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT for diagnosing a bacterial disease, as components of antibacterial
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT plants.
XX Example; SEQ ID NO 1311; 328pp; English.
PS The invention relates to isolated Acinetobacter baumannii nucleic acids.
XX The A. baumannii nucleic acids and polypeptides are useful as reagents
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of
CC A. baumannii and other Acinetobacter species in a sample, in screening
CC compounds for the ability to interfere with the A. baumannii life cycle
CC or to inhibit A. baumannii infection, and as biocontrol agents for
CC plants. The present sequence represents DNA encoding an A. baumannii
CC protein.
XX
SQ Sequence 375 BP; 122 A; 64 C; 93 G; 96 T; 0 U; 0 Other;
Query Match 14.6%; Score 85.8; DB 8; Length 375;
Best Local Similarity 57.4%; Pred. No. 1.1e-15;
Matches 175; Conservative 0; Mismatches 127; Indels 3; Gaps 1;
QY 235 GCAATTGTCTAGACCATGGAGATCCAGCAAGTTTTCATCGGCTTTACTGAAATCGGGATT 294
Db 52 GCAATTGTAAACCGTTTAAATTGGATGATGTGCGTGAAGCACTCTCTGACATTGGTGTGTA 111
QY 295 CGAGGTGTTACTGTTTCTGATGTGAGAGGTTTGGTGCAAGGAGGTTCTACCGAGAGA 354
Db 112 CAAGGGATTACCTTAACCTGAAGTTAAAGTTTGGTCGTACAAAGGACATACAGAACTT 171
QY 355 CACGTGGCTCTGAGTTCTCGGAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTT 414
Db 172 TACCGCGCGCTGAGTATGTGGTTGAT---TTCCTACCTAAAGTAAATCGAATTCG 228
QY 415 GTTAAGAAAGACCAAGTGGATCTGTATCAACACACAATAATTGAAGGACGACAGGACAGGA 474
Db 229 ATTAGTGATGAATGGTCGACGCGGTAATTGAGTCAATTACACGTGTGGCAAGCACTGGA 288
QY 475 GAGATTGGTGATGGCAAGATTTTGGTTTTCCTGTGTCAGATGTCATAAGAGTTAGGACA 534
Db 289 AAAATCGGCGACGGTAAGATTTTGTGACTAATCTGGAACAAGTCATCCGTATCCGTACA 348
QY 535 GGTGA 539
Db 349 GGTGA 353
RESULT 8
AAV21209_01
Continuation (2 of 17) of AAV21209 from base 100001 (Methanococcus jannaschii circular c
WP Sequence split into 17 fragments LOCUS AAV21209 Accession Aav21209
WP Fragment Name Begin End
WP AAV21209_00 1 110000
WP AAV21209_01 100001 210000

CC luminescens. Cells transformed to express the genes are useful for
CC recombinant production of the proteins, particularly toxins and
CC antibacterials useful as insecticides, bactericides and fungicides. The
CC genes, proteins, vectors containing the genes and Ab are also useful
CC therapeutically (to treat microbial infection by bacteria or fungi that
CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC biopesticides. Other uses of the genes and the proteins are as virulence
CC factors and for identifying targets of human diseases for which P.
CC luminescens is a model (particularly plague and whooping cough). This
CC sequence represents one of the isolated P. luminescens genes
XX

SQ Sequence 339 BP; 105 A; 51 C; 96 G; 87 T; 0 U; 0 Other;

Query Match 13.8%; Score 81.4; DB 7; Length 339;
Best Local Similarity 55.0%; Pred. No. 2.3e-14;
Matches 182; Conservative 0; Mismatches 146; Indels 3; Gaps 1;

QY	224	ACAAGTGGGAAGCAATTGTCAGACCATTGGAGAACATCCAGCAAGTTTCATCGGCTTTACTGA	283
Db	5	AAAAGATTGATCGGATTATATCAAAACCTTTCAAATTAGATGATGCGTGAAGCTCTGGCGG	64
QY	284	AAATCGGATTCGAGGTGTTACTGTTTCTGATGTGAGAGGGTTTGGTCACAAGGAGTT	343
Db	65	AAGTGGGTATCACCGGAATGACAGTACAGAGGTGAAGGTTTGGCGGCCAAAAGGTC	124
QY	344	CTACCGAGACACCGTGGCTCTGAGTTCTCGAAGACAAATTTGTTGTAAGTTAAGA	403
Db	125	ATACAGAGCTGTATCGCGGTGCAGATATATGTTGGAT--TTTCTGCCAAAAGTGAAA	181
QY	404	TGGAATCGTTGTTAAGAAAGACCAAGTGGAACTCTGTAATCAACACATAATTGAAGGAG	463
Db	182	TAGAAATTGTCGCCAGATGATATTGTCGATACCTGTGTTGAAACCATATGCAGACGG	241
QY	464	CAAGGACAGAGAGATTGGTGAATGGCAAGATTTTGTGTTTGCCTGTGTGATGTCATAA	523
Db	242	CACAGACCGGAAAATCGGTGATGGTAAATATTGTTATTTGATGTAGCACGTTGTGTGC	301
QY	524	GAGTTAGGACAGGTGAGCGTGGGGAGAAAGC	554
Db	302	GTATCCGCCACCGGTGACGAGGATGAAGAGGC	332

RESULT 11
ACF67367_52
Continuation (53 of 57) of ACF67367 from base 5200001 (Photorhabdus luminescens nucleotide
WP Sequence split into 57 fragments LOCUS ACF67367 Accession ACF67367

WP	Fragment Name	Begin	End
WP	ACF67367_30	1	110000
WP	ACF67367_31	100001	210000
WP	ACF67367_32	200001	310000
WP	ACF67367_33	300001	410000
WP	ACF67367_34	400001	510000
WP	ACF67367_35	500001	610000
WP	ACF67367_36	600001	710000
WP	ACF67367_37	700001	810000
WP	ACF67367_38	800001	910000
WP	ACF67367_39	900001	1010000
WP	ACF67367_40	1000001	1110000
WP	ACF67367_41	1100001	1210000
WP	ACF67367_42	1200001	1310000
WP	ACF67367_43	1300001	1410000
WP	ACF67367_44	1400001	1510000
WP	ACF67367_45	1500001	1610000
WP	ACF67367_46	1600001	1710000
WP	ACF67367_47	1700001	1810000
WP	ACF67367_48	1800001	1910000
WP	ACF67367_49	1900001	2010000
WP	ACF67367_50	2000001	2110000
WP	ACF67367_51	2100001	2210000
WP	ACF67367_52	2200001	2310000
WP	ACF67367_53	2300001	2410000
WP	ACF67367_54	2400001	2510000
WP	ACF67367_55	2500001	2610000

WP	ACF67367_26	2600001	2710000
WP	ACF67367_27	2700001	2810000
WP	ACF67367_28	2800001	2910000
WP	ACF67367_29	2900001	3010000
WP	ACF67367_30	3000001	3110000
WP	ACF67367_31	3100001	3210000
WP	ACF67367_32	3200001	3310000
WP	ACF67367_33	3300001	3410000
WP	ACF67367_34	3400001	3510000
WP	ACF67367_35	3500001	3610000
WP	ACF67367_36	3600001	3710000
WP	ACF67367_37	3700001	3810000
WP	ACF67367_38	3800001	3910000
WP	ACF67367_39	3900001	4010000
WP	ACF67367_40	4000001	4110000
WP	ACF67367_41	4100001	4210000
WP	ACF67367_42	4200001	4310000
WP	ACF67367_43	4300001	4410000
WP	ACF67367_44	4400001	4510000
WP	ACF67367_45	4500001	4610000
WP	ACF67367_46	4600001	4710000
WP	ACF67367_47	4700001	4810000
WP	ACF67367_48	4800001	4910000
WP	ACF67367_49	4900001	5010000
WP	ACF67367_50	5000001	5110000
WP	ACF67367_51	5100001	5210000
WP	ACF67367_52	5200001	5310000
WP	ACF67367_53	5300001	5410000
WP	ACF67367_54	5400001	5510000
WP	ACF67367_55	5500001	5610000
WP	ACF67367_56	5600001	5648894

Query Match 13.8%; Score 81.4; DB 7; Length 110000;

Best Local Similarity 55.0%; Pred. No. 2.8e-13;

Matches 182; Conservative 0; Mismatches 146; Indels 3; Gaps 1;

QY	224	ACAAGTGGGAAGCAATTGTCAGACCATTGGAGAACATCCAGCAAGTTTCATCGGCTTTACTGA	283
Db	49307	AAAAGATTGATCGGATTATCAAAACCTTTCAAATTAGATGATGCGTGAAGCTCTGGCGG	49366
QY	284	AAATCGGATTCGAGGTGTTACTGTTTCTGATGTGAGAGGGTTTGGTCACAAGGAGTT	343
Db	49367	AAGTGGGTATCACCGGAATGACAGTACAGAGGTGAAGGTTTGGCGGCCAAAAGGTC	49426
QY	344	CTACCGAGACACCGTGGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGA	403
Db	49427	ATACAGAGCTGTATCGCGGTGCAGATATATGTTGGAT---TTTCTGCCAAAAGTGAAA	49483
QY	404	TGGAATCGTTGTTAAGAAAGACCAAGTGGAACTCTGTAATCAACACATAATTGAAGGAG	463
Db	49484	TAGAAATTGTCGCCAGATGATATTGTCGATACCTGTGTTGAAACCATATATGCAGACGG	49543
QY	464	CAAGGACAGGAGAGATTGGTGAATGGCAAGATTTTGTGTTTGCCTGTGTGATGTCATAA	523
Db	49544	CACAGACCGGAAAATCGGTGATGGTAAATATTGTTATTTGATGTAGCACGTTGTGTGC	49603
QY	524	GAGTTAGGACAGGTGAGCGTGGGGAGAAAGC	554
Db	49604	GTATCCGCCACCGGTGACGAGGATGAAGAGGC	49634

RESULT 12

ACF65387_3

Continuation (4 of 7) of ACF65387 from base 300001 (Photorhabdus luminescens nucleotide
WP Sequence split into 7 fragments LOCUS ACF65387 Accession ACF65387

WP	Fragment Name	Begin	End
WP	ACF65387_0	1	110000
WP	ACF65387_1	100001	210000
WP	ACF65387_2	200001	310000
WP	ACF65387_3	300001	410000
WP	ACF65387_4	400001	510000
WP	ACF65387_5	500001	610000
WP	ACF65387_6	600001	696798

Query Match 13.8%; Score 81.4; DB 7; Length 110000;
Best Local Similarity 55.0%; Pred. No. 2.8e-13;
Matches 182; Conservative 0; Mismatches 146; Indels 3; Gaps 1;
QY 224 ACAAGGTGGAAGCAATTGTGACACCATGGAGAATCCAGCAAGTTTTCATCGGCTTTACTGA 283
Db 88821 AAAAGATTGATGCGATTATCAACCTTTCAAATTAGATGATGCGTGAAGCTCTGSCGG 88880
QY 284 AAATCGGGATTTCGAGGTTACTGTTTCTGATGTGAGAGGTTTGGTGCCACAAGGAGTT 343
Db 88881 AAGTGGGTATCACCGGATGACAGTAACAGAGGTGAAGGTTTGGCGGCCAAAAGGTC 88940
QY 344 CTACCGAGAGACACGGTGGCTCTGAGTTCTCGGAAGACAAATTTGTTGTTAAAGTTAAGA 403
Db 88941 ATACAGAGCTGTATCGGTCGAGAAATATATGTTGGTGGAT---TTCTGCCAAAAGTGAAA 88997
QY 404 TGGAAATCGTTGTTAAGAAAGACCAAGTGGAAATCTGTAATCAACACAATAATTGAAGGAG 463
Db 88998 TAGAAATTGTCGTGCCAGATGATATTGTGATACCTGTGTGAAACCAATTAATGCAGACGG 89057
QY 464 CAAGGACAGGAGAGATTGTTGATGGCAAGATTTTGTGTTTGCCTGTGTGAGATGTCATAA 523
Db 89058 CACAGACCGGGAAATCGGTGATGTTAAATATTTTGTATTGATGTAGCACGTGTTGTCG 89117
QY 524 GAGTTAGGACAGGTGAGCGTGGGAGAAAGC 554
Db 89118 GTATCCGCACCGGTGAGCAGGATGAAGAGGC 89148

RESULT 13
AAT42063_03
Continuation (4 of 19) of AAT42063 from base 300001 (Haemophilus influenzae complete genome)
WP Sequence split into 19 fragments LOCUS AAT42063 Accession Aat42063
WP Fragment Name Begin End
WP AAT42063_00 1 110000
WP AAT42063_01 100001 210000
WP AAT42063_02 200001 310000
WP AAT42063_03 300001 410000
WP AAT42063_04 400001 510000
WP AAT42063_05 500001 610000
WP AAT42063_06 600001 710000
WP AAT42063_07 700001 810000
WP AAT42063_08 800001 910000
WP AAT42063_09 900001 1010000
WP AAT42063_10 1000001 1110000
WP AAT42063_11 1100001 1210000
WP AAT42063_12 1200001 1310000
WP AAT42063_13 1300001 1410000
WP AAT42063_14 1400001 1510000
WP AAT42063_15 1500001 1610000
WP AAT42063_16 1600001 1710000
WP AAT42063_17 1700001 1810000
WP AAT42063_18 1800001 1830121

Query Match 13.6%; Score 79.8; DB 2; Length 110000;
Best Local Similarity 54.7%; Pred. No. 8.4e-13;
Matches 181; Conservative 0; Mismatches 147; Indels 3; Gaps 1;
QY 224 ACAAGGTGGAAGCAATTGTGACACCATGGAGAATCCAGCAAGTTTTCATCGGCTTTACTGA 283
Db 63919 AAAAAATCGAAGCAATGATTAAACCTTTAAATTAGACGATGTGCGAGAAAGTCTTTTCAG 63978
QY 284 AAATCGGGATTTCGAGGTTACTGTTTCTGATGTGAGAGGTTTGGTGCCACAAGGAGTT 343
Db 63979 ATATTGGTATTTCAGGTATGACAATCACAAGATACCGGATTTGTCGTCAAAAAGGTC 64038
QY 344 CTACCGAGAGACACGGTGGCTCTGAGTTCTCGGAAGACAAATTTGTTGTTAAAGTTAAGA 403
Db 64039 ATACAGAACTTTATCGTGTGCGGAATATATGTTGGTGGAT---TTTCTGCCGAAAGTGAAT 64095
QY 404 TGGAAATCGTTGTTAAGAAAGACCAAGTGGAAATCTGTAATCAACACAATAATTGAAGGAG 463

Db 64096 TCGAAGTGGTAGTTCTCTGATGAGCTTGTGGATCAATGTATTGAAGCATTAATTGAAACGG 64155
QY 464 CAAGGACAGGAGAGATTGTTGATGGCAAGATTTTGTGCTGTGTGATGTCAGATGTCAATA 523
Db 64156 CACAAACAGGTAAATCGGTGACGGCAAAATTTTGTGTTATCAGTTGAGAGAGCCATCC 64215
QY 524 GAGTTAGGACAGGTGAGCGTGGGAGAAAGC 554
Db 64216 GCATTGCCACGGCGGAAGAAACAGGATGC 64246

RESULT 14
ABZ40022
ID ABZ40022 standard; DNA; 363 BP.
XX
AC ABZ40022;
XX
DT 07-MAR-2003 (first entry)
XX
DE N. gonorrhoeae nucleotide sequence SEQ ID 4633.
XX
KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX
OS Neisseria gonorrhoeae.
XX
PN WO200279243-A2.
XX
PD 10-OCT-2002.
XX
PF 12-FEB-2002; 2002WO-IB002069.
XX
PR 12-FEB-2001; 2001GB-00003424.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Pontana MR, Pizza M, Massignani V, Monaci E;
XX
DR WPI; 2003-058415/05.
DR P-PSDB; ABP79052.
XX
PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
PT medicament for treating or preventing N. gonorrhoeae infection.
XX
PS Disclosure; Page 528; 815pp; English.
XX
CC The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
CC molecules of the invention
XX
SQ Sequence 363 BP; 84 A; 77 C; 123 G; 79 T; 0 U; 0 Other;

Query Match 13.1%; Score 77; DB 7; Length 363;
Best Local Similarity 54.1%; Pred. No. 5.3e-13;
Matches 180; Conservative 0; Mismatches 150; Indels 3; Gaps 1;
QY 224 ACAAGGTGGAAGCAATTGTGACACCATGGAGAATCCAGCAAGTTTTCATCGGCTTTACTGA 283
Db 32 AAAAAATCGAGGCGATTGTCAAACCGTTCAAGCTCGACGACGTGCGGAGGCGTTGACGG 91
QY 284 AAATCGGGATTTCGAGGTTACTGTTTCTGATGTGAGAGGTTTGGTGCCACAAGGAGTT 343
Db 92 AAATCGGCATTACGGCATGACCGTCAGCGAGGTCAAGGGTTCCGCGAGGAGAGGGGC 151
QY 344 CTACCGAGAGACACGGTGGCTCTGAGTTCTCGGAAGACAAATTTGTTGTTAAAGTTAAGA 403
Db 152 ATACGGAAATCTATCGCGGTGCGGAATATGCGCGTCGAT---TTCTGCCCAAGGTCAAAA 208
QY 404 TGGAAATCGTTGTTAAGAAAGACCAAGTGGAAATCTGTAATCAACACAATAATTGAAGGAG 463

Db	209	TCGAGTTGGTGTGGCGGATGATGCCGTGGACGCGGATTGACGTGATTGTCGAGGTGG	268
Qy	464	CAAGGACAGGAGAGATTGGTGATGGCAAGATTTTGTGTTTGCCTGTGTGATGTCTATAA	523
Db	269	CGCGTTCCGGGCAAAATCGGCGACGCGCAAGATTTTGTGCTGCCGGTCGAGGAGGCAATCC	328
Qy	524	GAGTTAGGACAGGTGAGCGTGGGGAGAGAAAGCAG	556
Db	329	GTATCCGACACGGGCGAACGTTCCGACGCGCGG	361
RESULT 15			
AAA81472			
ID	AAA81472 standard; DNA; 16526 BP.		
XX			
AC	AAA81472;		
XX			
DT	04-DEC-2000 (first entry)		
XX			
DE	N. meningitidis partial DNA sequence gnm_20 SEQ ID NO:20.		
XX			
KW	Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;		
KW	antigen; vaccine; diagnosis; infection; antibacterial; identification;		
KW	Meningococcus B; MenB; ds.		
XX			
OS	Neisseria meningitidis.		
XX			
PN	W0200022430-A2.		
XX			
PD	20-APR-2000.		
XX			
PF	08-OCT-1999; 99WO-US023573.		
XX			
PR	09-OCT-1998; 98US-0103794P.		
PR	30-APR-1999; 99US-0132068P.		
XX			
PA	(CHIR) CHIRON CORP.		
XX			
PI	Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;		
PI	Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;		
PI	Rappuoli R, Pizza M;		
XX			
DR	WPI; 2000-318079/27.		

	CC	that may be presumed targets for the immune system and which are not antigenically variable or at least more conserved than other more variable regions
SQ	Sequence	16526 BP; 3183 A; 4006 C; 5067 G; 4268 T; 0 U; 2 Other;
	Query Match	12.4%; Score 73; DB 3; Length 16526;
	Best Local Similarity	55.3%; Pred. No. 4.4e-11;
Matches	184; Conservative	0; Mismatches 145; Indels 4; Gaps 2;
QY	224	ACAAGGTGGAAGCAATTGTGCAGACCATGGAGAATCCACCGAAGTTTCATCGGCTTACTGA 283
Dd	13603	AAAAATCGAGGCGATTGTCAAACCGTTCAAATCTCGACGACGTGCGGAGGCCGTGACGG 13662
QY	284	AATCGGANTTCGAGGTGTTACTGTGTTTTCTGATGTGAGAGGGTTTGTCACACAGGAGTT 343
Dd	13663	AAATCGGCATTACGGGCATGACCGTCAGCGAGGTCAAAGGGTTCGGCAGGCAGAGGGGC 13722
QY	344	CTACCGAGAGACACGGTGGCTCTGAGTTCTCGGAAGACAAAATTTGTTGCTAAAGTTAAGA 403
Dd	13723	ATACGGAATCTATCGCGGCGCGGAATAACGCCGTGGAT---TTCCTGCCCAAAATCAAAA 13779
QY	404	TGGAATCGTTGTTAAGAAAGACCAAGTGGRAATCTGTAATCAACACAAATAATTGAAGGNG 463
Dd	13780	TOGAGCTGGTGTGGCGGATGATGCTGTGGRACGCCGGATTGACGTGAT-TTCGAGSGGG 13838
QY	464	CAAGGACAGGAGAGATTGGTGATCGCAAGATTTTTTGTTTTGTCCTGTGTCAGATGTCAFAA 523
Dd	13839	CSCGTTCCGGAAAAATCGGCGACGGCAAGATTTTTGTGTCGCCGGTTGAGGAGGUATCC 13898
QY	524	GAGTAGCACAGGTGAGCGTGGGGAGAAAGCAG 556
Dd	13899	GTATCCGACGCGCGAACGTTCCGACGCGGCAG 13931

Search completed: May 27, 2004, 19:59:42
Job time : 207.069 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 13:35:34 ; Search time 1755.75 Seconds
(without alignments)
14663.718 Million cell updates/sec

Title: US-09-756-541-16
Perfect score: 594
Sequence: 1 CGGCACGAGGCTACTGCGAA.....CTGACATGAGTACTTCTGCT 594

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl :

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_lr:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	594	100.0	594	6	AR125592	AR125592 Sequence
2	594	100.0	897	6	AR125590	AR125590 Sequence
3	592	99.7	840	8	AF095454	AF095454 Ricinus c
4	300.2	50.5	796	8	AY027892	AY027892 Medicago
5	285	48.0	588	6	AR125591	AR125591 Sequence
6	285	48.0	591	6	AX507438	AX507438 Sequence
7	285	48.0	591	8	BT005209	BT005209 Arabidops
8	285	48.0	817	6	AR125589	AR125589 Sequence
9	285	48.0	844	8	AF095455	AF095455 Arabidops
10	275.2	46.3	947	8	AY442185	AY442185 Lycopersi
11	216.2	36.4	902	8	AK068407	AK068407 Oryza sat
12	216.2	36.4	902	8	AK099152	AK099152 Oryza sat
13	215.4	36.3	1085	8	PPI489604	AJ489604 Pinus pin
14	107	18.0	686	1	AF017419	AF017419 Nostoc pu
15	103.2	17.4	2493	1	ASP251822	AJ251822 Anabaena
16	103.2	17.4	341880	1	AP003589	AP003589 Nostoc sp
17	99	16.7	384	1	FDGLNBPRT	X97327 F.diplosiph
18	98.8	16.6	191028	8	PPU38804	U38804 Porphyra pu
19	96.4	16.2	11978	1	U67574	U67574 Methanococc
20	96.4	16.2	110000	6	AR271569_01	Continuation (2 of
21	96.2	16.2	10338	1	U67464	U67464 Methanococc
22	96.2	16.2	110000	6	AR271569_14	Continuation (15 o
23	95.4	16.1	383	1	SYOGLNB	M62447 Synechococc
24	95.4	16.1	4348	1	AF079137	AF079137 Synechoco
25	94.4	15.9	452	1	SSGLNBP	X97496 Synechocyst
26	94.4	15.9	130001	1	D90915	D90915 Synechocyst
27	92	15.5	300143	1	AE017165	AE017165 Prochloro
28	90.4	15.2	164921	8	AF022186	AF022186 Cyanidium
29	90.4	15.2	299350	1	AP005370	AP005370 Thermosyn
30	88.4	14.9	9973	1	AE004295	AE004295 Vibrio ch
31	88.2	14.8	1622	1	AF120107	AF120107 Synechoco
32	86.8	14.6	300045	1	AE016803	AE016803 Vibrio vu
33	86.8	14.6	302320	1	AP005081	AP005081 Vibrio pa
34	86.4	14.5	262202	1	BX572094	BX572094 Prochloro
35	85.2	14.3	247950	1	AP005340	AP005340 Vibrio vu
36	84.8	14.3	96109	6	AR408756	AR408756 Sequence
37	84.8	14.3	96109	6	AX067460	AX067460 Sequence
38	84	14.1	11371	1	AE006236	AE006236 Pasteurel
39	83.6	14.1	375	6	AR318761	AR318761 Sequence
40	83.6	14.1	298950	1	AP004597	AP004597 Oceanobac
41	83.4	14.0	9977	1	AE002548	AE002548 Neisseria
42	83.2	14.0	300880	1	AE016917	AE016917 Chromobac
43	82.2	13.8	300272	1	AE017213	AE017213 Geobacter
44	82	13.8	346792	1	BX571658	BX571658 Wolinella
45	81.8	13.8	349061	1	NMA2Z2491	AL162753 Neisseria

ALIGNMENTS

RESULT 1
AR125592
LOCUS AR125592 594 bp DNA
DEFINITION Sequence 16 from patent US 6177275.
ACCESSION AR125592
VERSION AR125592.1 GI:14111654
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 594)
AUTHORS Coruzzi,G.M., Lam,H.-M. and Hsieh,M.-H.
TITLE Plant nitrogen regulatory P-PII genes
JOURNAL Patent: US 6177275-A 16 23-JAN-2001;
FEATURES Location/Qualifiers

source

1. .594

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match

Best Local Similarity

Matches 594; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

Qy

1

CGGCACGAGGCTACTCGGAAACTGGGCTTGCTCACTCCTCTTCATTTCTAATAACATCAAG

60

Db

1

CGGCACGAGGCTACTCGGAAACTGGGCTTGCTCACTCCTCTTCATTTCTAATAACATCAAG

60

Qy

61

AAAGAAATCCCTGTTTTCAGTTTGTTCAGAGCTTAGACATTCCTCGGTTT

120

Db

61

AAAGAAATCCCTGTTTTCAGTTTGTTCAGAGCTTAGACATTCCTCGGTTT

120

Qy

121

TCTCACTTTAACACCGCGGTCAAGCGGTAAAGATATGCCCGTGGTTCCTGTGATTAAT

180

Db

121

TCTCACTTTAACACCGCGGTCAAGCGGTAAAGATATGCCCGTGGTTCCTGTGATTAAT

180

Qy

181

GCCCAAAGCTCGCTGACTACATTCCTGATGCTAAATTCACAAAGTGAAGCAATTC

240

Db

181

GCCCAAAGCTCGCTGACTACATTCCTGATGCTAAATTCACAAAGTGAAGCAATTC

240

Qy

241

AGGCCCTGGCGAGTCTCGCAAGTTTCCTCGGCTTGTCTAAATAATTCGAGGTGT

300

Db

241

AGGCCCTGGCGAGTCTCGCAAGTTTCCTCGGCTTGTCTAAATAATTCGAGGTGT

300

Qy

301

ACTGTTCTGATGTTTCGAGGTTTGGTGCTCAAGTGGTTCAACTGAGAGCGAGGCGGC

360

Db

301

ACTGTTCTGATGTTTCGAGGTTTGGTGCTCAAGTGGTTCAACTGAGAGCGAGGCGGC

360

Qy

361

TCAGAAATTTCTGAAGACAAAGTTTGTGCTTAAAGTTAAGATGGAGATCGTGGTTAGCAA

420

Db

361

TCAGAAATTTCTGAAGACAAAGTTTGTGCTTAAAGTTAAGATGGAGATCGTGGTTAGCAA

420

Qy

421

GACCAAGTTGAGGATGTTATAGAAAATCAATTCAGGAGGCAAGAACTGGAGAGATTGGA

480

Db

421

GACCAAGTTGAGGATGTTATAGAAAATCAATTCAGGAGGCAAGAACTGGAGAGATTGGA

480

Qy

481

GACGGCAAGATTTCTTGCTGCCCTGTTTCAGATGTAATAAGAGTCCGCACTGGTGAGCGG

540

Db

481

GACGGCAAGATTTCTTGCTGCCCTGTTTCAGATGTAATAAGAGTCCGCACTGGTGAGCGG

540

Qy

541

GGTGATAAGGCTGAGAGGATGACAGGAGGGCGATCTGACATGAGTACTTCTGCT

594

Db

541

GGTGATAAGGCTGAGAGGATGACAGGAGGGCGATCTGACATGAGTACTTCTGCT

594

RESULT 2

AR125590

LOCUS

AR125590

897 bp

DNA

linear

PAT 16-MAY-2001

DEFINITION

Sequence 14 from patent US 6177275.

ACCESSION

AR125590

VERSION

AR125590.1

GI:14111652

KEYWORDS

Unknown.

SOURCE

Unknown.

ORGANISM

Unclassified.

REFERENCE

1 (bases 1 to 897)

AUTHORS

Coruzzi, G.M., Lam, H.-M. and Hsieh, M.-H.

TITLE

Plant nitrogen regulatory P-PII genes

JOURNAL

Patent: US 6177275-A 14 23-JAN-2001;

FEATURES

Location/Qualifiers

source

1. .897

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match

Best Local Similarity

Matches 594; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

Qy	1	CGGCACGAGGCTACTCGGA	AACTGGGCTTGCTCACTCCTCTTCATTTCTAATAACATCAAG	60
Db	50	CGGCACGAGGCTACTCGGA	AACTGGGCTTGCTCACTCCTCTTCATTTCTAATAACATCAAG	109
Qy	61	AAAGAAATCCCTGTTT	TGATTTTCAGTTTGTTCAGAGCTTAGACATTCCTCGGTTT	120
Db	110	AAAGAAATCCCTGTTT	TGATTTTCAGTTTGTTCAGAGCTTAGACATTCCTCGGTTT	169
Qy	121	TCTCACTTTAACACCGG	GTCAAGCGGTAAAGATATGCCCCCGTCTCCTGTGATTAAT	180
Db	170	TCTCACTTTAACACCGG	GTCAAGCGGTAAAGATATGCCCCCGTCTCCTGTGATTAAT	229
Qy	181	GCCCAAAGCTCGCTGAC	TACATTCCTGATGCTAAATTCACAAAGTGAAGCAATTC	240
Db	230	GCCCAAAGCTCGCTGAC	TACATTCCTGATGCTAAATTCACAAAGTGAAGCAATTC	289
Qy	241	AGGCCCTGCGGAGTCT	CGCAAGTTTCCTCGGCTTGTCTAAATAATTCGAGGTGT	300
Db	290	AGGCCCTGCGGAGTCT	CGCAAGTTTCCTCGGCTTGTCTAAATAATTCGAGGTGT	349
Qy	301	ACTGTTCTGATGTTTC	GAGGTTTGGTGCTCAAGTGGTTCAACTGAGAGCGAGGCGGC	360
Db	350	ACTGTTCTGATGTTTC	GAGGTTTGGTGCTCAAGTGGTTCAACTGAGAGCGAGGCGGC	409
Qy	361	TCAGAAATTTCTGAAG	CAAGTTTGTCTGCTTAAAGTTAAGATGGAGATCGTGGTTAGCAA	420
Db	410	TCAGAAATTTCTGAAG	CAAGTTTGTCTGCTTAAAGTTAAGATGGAGATCGTGGTTAGCAA	469
Qy	421	GACCAAGTTGAGGATG	TTATAGAAAATCAATTCAGGAGGCAAGAACTGGAGAGATTGGA	480
Db	470	GACCAAGTTGAGGATG	TTATAGAAAATCAATTCAGGAGGCAAGAACTGGAGAGATTGGA	529
Qy	481	GACGGCAAGATTTCTT	GTGCTGCCCTGTTTCAGATGTAATAAGAGTCCGCACTGGTGAGCGG	540
Db	530	GACGGCAAGATTTCTT	GTGCTGCCCTGTTTCAGATGTAATAAGAGTCCGCACTGGTGAGCGG	589
Qy	541	GGTGATAAGGCTGAG	AGGATGACAGAGGGCGATCTGACATGAGTACTTCTGCT	594
Db	590	GGTGATAAGGCTGAG	AGGATGACAGAGGGCGATCTGACATGAGTACTTCTGCT	643
RESULT 3				
AF095454				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
1. .840				
/organism="Ricinus communis"				
/mol_type="mRNA"				
/db_xref="taxon:3988"				
<1. .595				
CDS				

840 bp mRNA linear PLN 19-NOV-1998
AF095454
Ricinun communis PII protein mRNA, partial cds.
AF095454
AF095454.1 GI:3885940
Ricinun communis (castor bean)
Ricinun communis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Malpighiales; Euphorbiaceae; Acalyphoideae; Acalyphaeae; Ricinus.
1 (bases 1 to 840)
Hsieh, M.H., Lam, H.M., van de Loo, F.J. and Coruzzi, G.
A PII-like protein in Arabidopsis: putative role in nitrogen sensing
Proc. Natl. Acad. Sci. U.S.A. 95 (23), 13965-13970 (1998)
99030678
9811909
2 (bases 1 to 840)
Hsieh, M.H., Lam, H.M., van de Loo, F.J. and Coruzzi, G.M.
Direct Submission
Submitted (29-SEP-1998) Biology, New York University, 100 Washington Square East, New York, NY 10003, USA
Location/Qualifiers
1. .840
/organism="Ricinus communis"
/mol_type="mRNA"
/db_xref="taxon:3988"
<1. .595

AUTHORS Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Carninci,P., Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Hsuan,V.W., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.

TITLE Arabidopsis ORF clones

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 591)

AUTHORS Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Carninci,P., Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Hsuan,V.W., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.

TITLE Direct Submission

JOURNAL Submitted (14-MAR-2003) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGECC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Davis,R.W., Theologis,A., and Ecker,J.R.

Kim,C.J. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES source Location/Qualifiers 1..591 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /chromosome="4" /clone="U23463" /ecotype="Columbia" /note="This clone is in pUNI 51" 1..591 /note="P II nitrogen sensing protein GLB I" /codon_start=1 /product="At4g01900" /protein_id="AA063273.1" /db_xref="GI:28950699" /translation="MAASMTKPISTISLGFYSDRKNIAFSDCISCSGFRHSRPSCLD LVTKSPNSNRVLPVVSAGQISSDIYIPDSKFYKVEAIVRFWRIOQVSSALLKIGIRGVT VSDVRGFGAQQGSTERHGSSEFDKPFVAKVMEIVVKDQVESVINTIIEGARTGEI GDGKIFVLPVSDVIRVTRTGERGEKAEKMTGDMLSPS"

CDS 1..591 /note="P II nitrogen sensing protein GLB I" /product="At4g01900" /protein_id="AA063273.1" /db_xref="GI:28950699" /translation="MAASMTKPISTISLGFYSDRKNIAFSDCISCSGFRHSRPSCLD LVTKSPNSNRVLPVVSAGQISSDIYIPDSKFYKVEAIVRFWRIOQVSSALLKIGIRGVT VSDVRGFGAQQGSTERHGSSEFDKPFVAKVMEIVVKDQVESVINTIIEGARTGEI GDGKIFVLPVSDVIRVTRTGERGEKAEKMTGDMLSPS"

ORIGIN Query Match 48.0%; Score 285; DB 8; Length 591; Best Local Similarity 70.3%; Pred. No. 4.5e-68; Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1; QY 7 GAGGCTACTGCGAAACTGGGCTTGCTCACTCTCTTCAATCTAATAACATCAAGAAGAA 66 Db 7 GCGTCAATGACGAAACCCATCTCAATAACTCTCTCGGTTTCTATTCTGATCGAAGAAC 66 QY 67 TTCCCTGTTTTGATTTTCAGTTTGTGTTTGTCCAGAGCTTAGACATTTCTCGGTTTCTCAC 126

Db 67 ATTGCTTTCTCTGATTCGATTTCGATTGTTTCTGGATTTCAGACATTCCTCCGACCATCTGC 126 QY 127 TTTAACACACCGCGTCAAGCGCGTAAGATATGCC---CCCCTGTTCTCTGATTAATGCC 183 Db 127 CTCGATTGTTGTCACAAAGTCACCGAGTAAACAGTCGTGTTTACCTGTCGTTAGTGCC 186 QY 184 CAAAGCTCGCTGACTACATTCCTGATGCTAAATTTCTACAAAGTGAAGCAATTTCTCAGG 243 Db 187 CAAATATCTTCTGATTATATTCAGACTCGAAATTTTCAAGGTGAAGCAATTTGTGAGA 246 QY 244 CCCTGGCGAGTCTCGCAAGTTTCTCGGCTTCTGCTTAAATAATTTGTTATTCGAGGTGTACT 303 Db 247 CCATGGAGATCCAGCAAGTTTTCATCGGCTTTACTGAAAATCGGATTCGAGGTGTACT 306 QY 304 GTTCTCTGATGTTTCGAGGTTTTCGTTGCTCAAGTGGTTCAACTGAGAGCGGCGGCTCA 363 Db 307 GTTCTCTGATGAGAGGGTTTGGTGCAACAAGGAGGTTCTACCGAGAGACACGCTGCTCT 366 QY 364 GAATTTTCTGAGACAAAGTTTGTGTTGCTAAAGTTAAGATGGAGATCGTGGTTAGCAAGAC 423 Db 367 GAGTTCTCGGAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAGAAAGAC 426 QY 424 CAGTTGAGGATGTTATAGAAAAATCAATTGAGGAGGCAAGAACTGGAGAGATTGGAGAC 483 Db 427 CAGTGAATCTGTAATCAACACAAATATTGAAGGACCAAGGACAGGAGAGATTGGTAT 486 QY 484 GGCAGATTTTCTGCTGCTGCTGTTTTCAGATGTAATAAGAGTCCGCACTGGTGAGCGGGT 543 Db 487 GGCAGATTTTGTGTTGCTGTTGTCAGATGTCATAAGAGTTAGGACAGGTGAGCGTGGG 546 QY 544 GATAAGGCTGAGAGGATGACAGGAG 568 Db 547 GAGAAAGCAGAGAGAGATGACTGGTG 571

RESULT 8

AR125589

LOCUS AR125589 817 bp DNA linear PAT 16-MAY-2001

DEFINITION Sequence 13 from patent US 6177275.

ACCESSION AR125589

VERSION AR125589.1 GI:14111651

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 817)

AUTHORS Coruzzi,G.M., Lam,H.-M. and Hsieh,M.-H.

TITLE Plant nitrogen regulatory P-PII genes

JOURNAL Patent: US 6177275-A 13 23-JAN-2001;

FEATURES source Location/Qualifiers 1..817 /organism="unknown" /mol_type="unassigned DNA"

ORIGIN Query Match 48.0%; Score 285; DB 6; Length 817; Best Local Similarity 70.3%; Pred. No. 4.7e-68; Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1; QY 7 GAGGCTACTGCGAAACTGGGCTTGCTCACTCTCTTCAATCTAATAACATCAAGAAGAA 66 Db 39 GCGTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGAAGAAC 98 QY 67 TTCCCTGTTTTGATTTTCAGTTTGTGTTTGTCCAGAGCTTAGACATTTCTCGGTTTCTCAC 126 Db 99 ATTGCTTTCTCTGATTCGATTTCGATTGTTGTTCTGGATTTCAGACATTCCTCCGACCATCTGC 158 QY 127 TTTAACACCGCGTCAAGCGCGTAAAGATATGCC---CCCCTGTTCTCTGTTGATTAATGCC 183 Db 159 CTCGATTGTTGGTCACAAAGTCACCGAGTAAATAACAGTCGTGTTTACCTGTCGTTAGTGCC 218 QY 184 CAAAGCTCGCTGACTACATTTCTGATGCTAATAATTTCTAAGTGAAGCAATTTCTCAGG 243

Db 219 CAAATATCTTCTGATTATATCCAGACTCGAAATTTTACAAGGTGGAAGCAATTTGTCAGA 278

Qy 244 CCCTGGGAGTCTCGCAAGTTTCTCGGCTTTGCTAAATAATTTGGTATTCGAGGTGTTACT 303

Db 279 CCATGGAGATCCAGCAAGTTTTCATCGGCTTTACTGAAATCGGATTCGAGGTGTTACT 338

Qy 304 GTTCTCTGATGTTTCGAGGTTTGGTGCTCAAGTGTTCAACTGAGAGGCGGGGCTCA 363

Db 339 GTTCTCTGATGTGAGAGGTTTGGTGCAAGGAGGTCTTACCGAGAGACACCGTGGCTCT 398

Qy 364 GAAATTTCTGAAGACAAAGTTTGTGCTAAAGTTAAGATCGGAGATCGTGTAGCAAGAC 423

Db 399 GAGTTCGGAAGACAAATTTGTGCTAAAGTTAAGATCGGAAATCGTTGTTAAGAAAGAC 458

Qy 424 CAGGTTGAGGATGTTATAGAAAAAATCATTTGAGGAGCGCAAGAACTGGAGAGATTGGAGAC 483

Db 459 CAAGTGGATCTGTAATCAACACAATAATTGAAGGAGCAAGGACAGGAGATTGGTGAT 518

Qy 484 GGCAAGATTTTCTTGCTGCTGTTTCAGATGTAATAAGAGTCCGCACTGGTGAGCGGGT 543

Db 519 GGCAAGATTTTGTGTTTGGCTGTGTGATGTCATAGAGTTAGGACAGGTGAGCGTGGG 578

Qy 544 GATAAGGCTGAGAGGATGACAGGAG 568

Db 579 GAGAAAGCAGAGAAGATGACTGGTG 603

RESULT 9

AF095455

LOCUS

DEFINITION Arabidopsis thaliana PII protein mRNA, complete cds.

ACCESSION AF095455

VERSION AF095455.1 GI:3885942

KEYWORDS

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 844)

AUTHORS Hsieh,M.H., Lam,H.M., van de Loo,F.J. and Coruzzi,G.

TITLE A PII-like protein in Arabidopsis: putative role in nitrogen sensing

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (23), 13965-13970 (1998)

MEDLINE 99030678

PUBMED 9811909

REFERENCE 2 (bases 1 to 844)

AUTHORS Hsieh,M.H., Lam,H.M., van de Loo,F.J. and Coruzzi,G.M.

TITLE Direct Submission

JOURNAL Submitted (29-SEP-1998) Biology, New York University, 100 Washington Square East, New York, NY 10003, USA

FEATURES

source

1. .844

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="IV"

/map="10.8 cm"

31. .621

/function="putative role in nitrogen sensing"

/note="GLB1"

/codon_start=1

/product="PII protein"

/protein_id="AAC78333.1"

/db_xref="GI:3885943"

/translation="MAASMTKPIISITSLGFYSDRKNIAFSDCISICSGFRHSRPSCLD

LVTKSPNSNRVLPVVSQAISDDYIPDSKFYKVEAIVRPWRIQQVSSALLKIGIRGVT

VSDVRGFGAQQGSTERHGGSEPSSEDKFVAKVYMEIVVKDKQVESVINTIIGARTGEI

GDGKIFVLPVSDVIRVIRGTGERGEKAERMTGDMLSPS"

ORIGIN

Query Match 48.0%; Score 285; DB 8; Length 844;

Best Local Similarity 70.3%; Pred. No. 4.7e-68;

Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

Qy 7 GAGGCTACTGCGAAACTGGCTTGCTCACTCCTCTTCTATCTAATAACATCAAGAAAGAA 66

Db 37 GCGTCAATGACGAAACCCCATCTCAATAACTTCTCTCGGTTTCTATTCTGTATCGAAAGAAC 96

Qy 67 TTCCCTGTTTTTGATTTTCAGTTTGTGTTTGTCCAGAGCTTTAGACATTTCTCGGTTTCTCAC 126

Db 97 ATTGCTTTCTCTGATTGCAATTCGATTGTTCTGGATTGAGACATTCGACCATCTTGC 156

Qy 127 TTTAACACCGCGGTCAAGCGGTAAGATATGCC---CCCGTCGTTCTGTGATTAATGCC 183

Db 157 CTCGATTTGGTCACAAAGTCACCGAGTAATAACAGTCGTGTTTACCTGTCTGTAGTGC 216

Qy 184 CAAAGCTCGCTGACTACATTCCTGATGCTAAATTTCTACAAAGTGAAGCAATTTCTCAGG 243

Db 217 CAAATATCTTCTGATTATATTCAGACTCTCGAAATTTTACAGGTGGAAGCAATTTGTGAGA 276

Qy 244 CCCTGGCGAGTCTCGCAAGTTTCTCGGCTTCAAGGTGGTTCAACTGAGAGGCAAGCGGCTCA 303

Db 277 CCATGGAGATCCAGCAAGTTTTCATCGGCTTTACTGAAATCGGGATTGAGGTGTTACT 336

Qy 304 GTTCTGATGTTTCGAGGTTTTTGTGCTCAAGGTGGTTCAACTGAGAGGCAAGCGGCTCA 363

Db 337 GTTCTGATGTGAGAGGTTTGTGTCACAAGGAGGTTCTACCGAGAGACACGGTGGCTCT 396

Qy 364 GAATTTCTGAAGACAAAGTTTGTGCTAAAGTTAAGATGGAGATCGTGGTTAGCAAAGAC 423

Db 397 GAGTCTCGGAAGACAAATTTGTGCTAAAGTTAAGATGGAATCGTTGTTAAGAAAGAC 456

Qy 424 CAGGTTGAGGATGTTATAGAAAAATCATTTGAGGAGGCAAGACTGGAGAGATTGGAGAC 483

Db 457 CAAGTGAATCTGTAATCAACACAATAATTGAAGGAGCAAGGACAGGAGAGATTGGTGAT 516

Qy 484 GGCAAGATTTTCTTGCTGCTGTTTCAGATGTAATAAGATCCGCACTGGTGAGCGGGT 543

Db 517 GGCAAGATTTTGTGTTTGGCTGTGTCAGATGTATAAGATTAGCACAGGTGAGCGTGGG 576

Qy 544 GATAAGGCTGAGAGGATGACAGGAG 568

Db 577 GAGAAAGCAGAGAAGATGACTGGTG 601

RESULT 10

AY442185

LOCUS

DEFINITION Lycopersicon esculentum PII-like protein (GLB1) mRNA, complete cds.

ACCESSION AY442185

VERSION AY442185.1 GI:38231569

KEYWORDS

SOURCE Lycopersicon esculentum (tomato)

ORGANISM Lycopersicon esculentum

REFERENCE 1 (bases 1 to 947)

AUTHORS Walch-Liu,P., Roemheld,V. and von Wiren,N.

TITLE A PII-like protein in tomato

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 947)

AUTHORS Walch-Liu,P., Roemheld,V. and von Wiren,N.

TITLE Direct Submission

JOURNAL Submitted (17-OCT-2003) Institute of Plant Nutrition, Hohenheim University, Fruwirthstr.20, Stuttgart 70593, Germany

FEATURES

source

1. .947

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/db_xref="taxon:4081"

1. .947

/gene="GLB1"

/note="LeGLB1"

53. .652

CDS

QY	249	GC	GAGTCTCGCAAGTTTCCTCGGCTTTGCTAAAAATTGGTATT	CGAGGTGTTACTGTTTC	308
DB	327	GAGGGTGCCTTATGTGTCATCGGGTTTGCTGCAATGSGGATCAGAGGCGTGACGGTGTC	386		
QY	309	TGATGTTTCGAGGTTTTGGTGCTCAAGGTGGTTCAACTGAGAGGCAGGGCGGCTCAGAAATT	368		
DB	387	CGACGTGCGGGGTTTCGGCGCACAGGGCGGGTCAACTGAGAGGCATGAAGGGTCAGAAATT	446		
QY	369	TTCTGAAGACAAGTTTTGTGTTAAAGTTAAGATCGGAGATCGTGGTTAGCAAGACCCAGGT	428		
DB	447	TGCAGAAGATACATTTTATTGATAAAGTTAAGATGGAATAAGTGGTGTCCAAGGATCAGGT	506		
QY	429	TGAGGATGTTTATAGAAAAATCATTTGAGGAGGCCAAGAACTGGAGAGATTGGAGACGGCAA	488		
DB	507	TGAAGCTGTTGTTGACAAAGATAAATTGAAAAGGCCAAGAACAGGAGAAATTGGTGTGGAAA	566		
QY	489	GATTTTCTCGCTGCCGTGTTTCAGATGTAATAAGAGTCCGCACTGGTGAGCGGGGTGATAA	548		
DB	567	AATATTTTGTATACCGGTGTGGACGTGATCAGAATACGCACCGCGCAACGAGGGGAGCG	626		
QY	549	GGCTGAGAGGATGACAGGAGGGCGATCTGCATGAGTACTTCTGC	593		
DB	627	AGCGGAGAGGATGGCCGAGGGCTGGCGGACAAGCTGTCTCTCAGC	671		

ORIGIN

Query Match	36.3%	Score 215.4;	DB 8;	Length 1085;
Best Local Similarity	73.2%	Pred. No. 1e-48;		
Matches 276;	Conservative	0;	Mismatches 101;	Indels 0;
			Gaps	0;

QY	193	CCTGACTACATTCTCTGATGCTAAATCTACAAAGTGAAGCAATCTCAGGCCCTGGCGA	252
Db	327	CCAGATTANGTCCAGAAAGCCAACTTTTACAAAGTAGAAGCAATATTGAGGCCATGGCGC	386
QY	253	GTCTCGCAAGTTTCCTGGCTTTGCTAAATAATGGTATTTCGAGGTGTTACTGTTCCTGAT	312
Db	387	ATCTCCCATGTGACTACGGGTCTATTGAAAATGGGGATTTCGTGGCGTAACCTGTCCTGAT	446
QY	313	GTTCGAGGTTTTGGTGCTCAAGGTGTTCAACTGAGAGGCGAGGGCGCTCAGAAATTTCT	372
Db	447	GTTAGAGGTTTTGGAGTTCAGGCTGGATCTGCAGAACGGCAAGCAGGCTCCGAGTTTCT	506
QY	373	GAAGACAAGTTTGTCTGCTAAAGTTAAGATGGAGATCGTGTTAGCAAGACAGGTTGAG	432
Db	507	AAAGACAATTTGTGTCAAAAGTAAAGATGGAGATTGTTGTTATCTAAAGATCAGGTAGAA	566
QY	433	GATGTTATAGAAAAAATCATTGAGGAGGCGCAAGAACTGGACAGATTGGAGACGGCAAGATT	492
Db	567	GCAGTAATTGATGCAATCATTTGATGAGGCGCAAGAACTGGAGAAATTTGGAGATGGCAATA	626
QY	493	TTCTTGCTGCCGTGTTTCAGATGTAAATAAGATCCGCACTGGTGAGCGGGGTGATAAGGCT	552
Db	627	TTTGTGGTTCCAGTTGCAGATGTCAATTCGTGTGAGAACAGGTGAGCGGGGACTTGAAGCCA	686
QY	553	GAGAGGATGACAGGAGG	569
Db	687	GAGAGAATGGCTGGGTGG	703

ORIGIN

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 13:28:59 ; Search time 207.162 Seconds
(without alignments)
12180.972 Million cell updates/sec

Title: US-09-756-541-16
Perfect score: 594
Sequence: 1 CGGCACGAGCTACTGGAA.....CTGACATGAGTACTTCTGCT 594

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	594	100.0	594	AAF58584	Aaf58584 Ricinus c
2	594	100.0	897	AAF58582	Aaf58582 Ricinus c
3	285	48.0	588	AAF58583	Aaf58583 Arabidops
4	285	48.0	591	ABZ14328	Abz14328 Arabidops
5	285	48.0	817	AAF58581	Aaf58581 Arabidops
6	96.4	16.2	110000	2 AAV21209_01	Continuation (2 of
7	96.2	16.2	110000	2 AAV21209_14	Continuation (15 o
8	84.8	14.3	96109	4 AAF28548	Aaf28548 Genomic f
9	83.6	14.1	375	8 ADA30024	Ada30024 DNA encod
10	81.4	13.7	339	7 ACF72027	Acf72027 Photorhab
11	81.4	13.7	110000	7 ACF67367_52	Continuation (53 o
12	81.4	13.7	110000	7 ACF65387_3	Continuation (4 of
13	77.6	13.1	110000	2 AAT42063_03	Continuation (4 of
14	77	13.0	363	7 ABZ40022	Abz40022 N. gonorr
15	73	12.3	16526	3 AAA81472	Aaa81472 N. mening
16	73	12.3	110000	3 AAA81489_6	Continuation (7 of
17	73	12.3	172325	3 AAF21613	Aaf21613 Neisseria
18	73	12.3	349980	3 AAF21612	Aaf21612 Neisseria
19	70.8	11.9	336	5 AAH68445	Aah68445 C glutami
20	70.8	11.9	417	5 AAH67222	Aah67222 C glutami
21	70.8	11.9	349980	5 AAH68531	Aah68531 C glutami
22	64.8	10.9	290	4 AAF71296	Aaf71296 Corynebac
23	64.2	10.8	1164	5 AAS88441	Aas88441 DNA encod

24	64.2	10.8	11095	4 AAS46256	Aas46256 DNA encod
25	60.4	10.2	110000	6 ABA90521_16	Continuation (17 o
26	56.2	9.5	110000	6 ABA03041_15	Continuation (16 o
27	55.2	9.3	339	7 ACF71482	Acf71482 Photorhab
28	55.2	9.3	663	7 ACF65588	Acf65588 Photorhab
29	55.2	9.3	110000	7 ACF67367_46	Continuation (47 o
30	55.2	9.3	110000	7 ACF65388_01	Continuation (2 of
31	54.8	9.2	110000	4 AAI99682_32	Continuation (33 o
32	54.8	9.2	110000	4 AAI99683_32	Continuation (33 o
33	52.6	8.9	174	7 ABZ39932	Abz39932 N. gonorr
34	52.2	8.8	2751	6 ABQ70750	Abq70750 Listeria
35	50.8	8.6	1970	3 AAZ99179	Aaz99179 T. hydrot
36	49	8.2	110000	6 ABQ69245_15	Continuation (16 o
37	49	8.2	110000	6 ABQ67195_1	Continuation (2 of
38	49	8.2	110000	6 ABQ67195_2	Continuation (3 of
39	48	8.1	862	6 ABQ68551	Abq68551 Listeria
40	45.4	7.6	342	6 ABN18985	Abn18985 Human ORF
41	41.4	7.0	674	7 ACF67052	Acf67052 Photorhab
42	38.8	6.5	1089	7 ACA44542	Aca44542 Prokaryot
43	38.4	6.5	2208	5 AAS88439	Aas88439 DNA encod
44	37.8	6.4	3858	5 AAS89065	Aas89065 DNA encod
45	37.8	6.4	3858	5 AAS90406	Aas90406 DNA encod

ALIGNMENTS

RESULT 1

AAF58584

ID AAF58584 standard; cDNA; 594 BP.

XX AAF58584;

XX 23-APR-2001 (first entry)

XX Ricinus communis P-PII cDNA fragment.

XX Castor bean; PII; plant nitrogen regulatory gene; P-PII;

XX nitrogen assimilation; transgenic plant; herbicide screening; ss.

XX Ricinus communis.

XX US6177275-B1.

XX 23-JAN-2001.

XX 23-JUL-1997; 97US-00899330.

XX 24-JUL-1996; 96US-0022328P.

XX (UUNY) UNIV NEW YORK STATE.

XX Coruzzi GM, Lam H, Hsieh M;

XX WPI; 2001-158572/16.

XX P-PSDB; AAB69496.

XX Novel P-PII genes capable of regulating plant nitrogen assimilation, useful for transgenic plant production, and as probes for isolating additional genomic clones having P-PII gene promoters.

XX Claim 2; Col 37-38; 35pp; English.

XX The present sequence encodes a nitrogen regulatory PII protein. Novel plant PII (also called P-PII) nucleotide sequences have been isolated. They are useful for regulating nitrogen assimilation in plants, and in transgenic plant production. They are also used to engineer organisms that overexpress wild-type or mutant P-PII regulatory proteins. P-PII proteins are useful for in vitro screening of herbicides. P-PII nucleotides may be used as probes for isolating additional genomic clones with the promoters of P-PII genes. P-PII promoters are light- and/or sucrose-inducible, and are suitable for genetic engineering of plants

SQ Sequence 594 BP; 149 A; 117 C; 157 G; 171 T; 0 U; 0 Other;
Query Match 100.0%; Score 594; DB 4; Length 594;
Best Local Similarity 100.0%; Pred. No. 2.1e-168;
Matches 594; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGCACGAGCTACTGCGAAACTGGGCTTGCTCACTCTCTTCAATCTAATAACATCAAG 60
Db 1 CGGCACGAGCTACTGCGAAACTGGGCTTGCTCACTCTCTTCAATCTAATAACATCAAG 60
QY 61 AAAGAAATCCCTGTTTTTGATTTTCAGTTTGTGTCAGAGCTTAGACATTCGGTTT 120
Db 61 AAAGAAATCCCTGTTTTTGATTTTCAGTTTGTGTCAGAGCTTAGACATTCGGTTT 120
QY 121 TCTCACTTTAAACACCGCGGTCAAGCGCTAAGATATGCCCGCTCGTTCCTGTGATTAAT 180
Db 121 TCTCACTTTAAACACCGCGGTCAAGCGCTAAGATATGCCCGCTCGTTCCTGTGATTAAT 180
QY 181 GCGCAAAGCTCGCCTGACTACATTCCTGTGCTAAATTTCTACAAAGTGAAGCAATCTC 240
Db 181 GCGCAAAGCTCGCCTGACTACATTCCTGTGCTAAATTTCTACAAAGTGAAGCAATCTC 240
QY 241 AGGCCCTGGCAGTCTCGCAAGTTTCCTCGGCTTTGCTAAATAATTGGTATTCGAGTGT 300
Db 241 AGGCCCTGGCAGTCTCGCAAGTTTCCTCGGCTTTGCTAAATAATTGGTATTCGAGTGT 300
QY 301 ACTGTTTCTGATGTTTCGAGGTTTTCGAGGTTTTCGAGGTTTTCGAGGTTTTCGAGG 360
Db 301 ACTGTTTCTGATGTTTCGAGGTTTTCGAGGTTTTCGAGGTTTTCGAGGTTTTCGAGG 360
QY 361 TCAGAAATTTCTGAAGACAAGTTTGTGCTAAAGTTAAGATGAGATCGTGGTTAGCAAA 420
Db 361 TCAGAAATTTCTGAAGACAAGTTTGTGCTAAAGTTAAGATGAGATCGTGGTTAGCAAA 420
QY 421 GACCAGGTTGAGGATGTTATAGAAAAATCATTTGAGGAGGCAAGAACTCGAGAGATTGGA 480
Db 421 GACCAGGTTGAGGATGTTATAGAAAAATCATTTGAGGAGGCAAGAACTCGAGAGATTGGA 480
QY 481 GACGGCAAGATTTCTTGCTGCCTGTTTCAGATGTAATAAGATCCGCACTGGTGAGCGG 540
Db 481 GACGGCAAGATTTCTTGCTGCCTGTTTCAGATGTAATAAGATCCGCACTGGTGAGCGG 540
QY 541 GGTGATAAGCTGAGAGGATGACAGGAGGCGATCTGACATGAGTACTTCTGCT 594
Db 541 GGTGATAAGCTGAGAGGATGACAGGAGGCGATCTGACATGAGTACTTCTGCT 594

RESULT 2
AAF58582
ID AAF58582 standard; cDNA; 897 BP.
XX
AC AAF58582;
XX
DT 23-APR-2001 (first entry)
XX
DE Ricinus communis P-II cdNA.
XX
KW Castor bean; PII; plant nitrogen regulatory gene; P-II;
nitrogen assimilation; transgenic plant; herbicide screening; ss.
XX
OS Ricinus communis.
XX
PN US6177275-B1.
XX
PD 23-JAN-2001.
XX
PF 23-JUL-1997; 97US-00899330.
PR 24-JUL-1996; 96US-0022328P.
XX
PA (UYNV) UNIV NEW YORK STATE.
XX
PI Coruzzi GM, Lam H, Hsieh M;

XX WPI; 2001-158572/16.
DR P-PSDB; AAB69496.
XX
PT Novel P-II genes capable of regulating plant nitrogen assimilation,
useful for transgenic plant production, and as probes for isolating
additional genomic clones having P-II gene promoters.
XX
PS Claim 2; Fig 13; 35pp; English.
XX
CC The present sequence encodes a nitrogen regulatory PII protein. Novel
plant PII (also called P-II) nucleotide sequences have been isolated.
CC They are useful for regulating nitrogen assimilation in plants, and in
transgenic plant production. They are also used to engineer organisms
CC that overexpress wild-type or mutant P-II regulatory proteins. P-II
proteins are useful for in vitro screening of herbicides. P-II
CC nucleotides may be used as probes for isolating additional genomic clones
CC with the promoters of P-II genes. P-II promoters are light- and/or
sucrose-inducible, and are suitable for genetic engineering of plants
XX
SQ Sequence 897 BP; 254 A; 171 C; 218 G; 254 T; 0 U; 0 Other;
Query Match 100.0%; Score 594; DB 4; Length 897;
Best Local Similarity 100.0%; Pred. No. 2.5e-168;
Matches 594; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGCACGAGGCTACTGCGAAACTGGGCTTGCTCACTCTCTTCAATCTAATAACATCAAG 60
Db 50 CGGCACGAGGCTACTGCGAAACTGGGCTTGCTCACTCTCTTCAATCTAATAACATCAAG 109
QY 61 AAAGAAATCCCTGTTTTTGATTTTCAGTTTGTGTCAGAGCTTAGACATTCGGTTT 120
Db 110 AAAGAAATCCCTGTTTTTGATTTTCAGTTTGTGTCAGAGCTTAGACATTCGGTTT 169
QY 121 TCTCACTTTAAACACCGCGGTCAAGCGCGTAAGATATGCCCGCTCGTTCCTGTGATTAAT 180
Db 170 TCTCACTTTAAACACCGCGGTCAAGCGCGTAAGATATGCCCGCTCGTTCCTGTGATTAAT 229
QY 181 GCGCAAAGCTCGCCTGACTACATTCCTGTGCTAAATTTCTACAAAGTGAAGCAATCTC 240
Db 230 GCGCAAAGCTCGCCTGACTACATTCCTGTGCTAAATTTCTACAAAGTGAAGCAATCTC 289
QY 241 AGGCCCTGGCAGTCTCGCAAGTTTCCTCGGCTTTGCTAAATAATTGGTATTCGAGTGT 300
Db 290 AGGCCCTGGCAGTCTCGCAAGTTTCCTCGGCTTTGCTAAATAATTGGTATTCGAGTGT 349
QY 301 ACTGTTTCTGATGTTTCGAGGTTTTCGAGGTTTTCGAGGTTTTCGAGGTTTTCGAGG 360
Db 350 ACTGTTTCTGATGTTTCGAGGTTTTCGAGGTTTTCGAGGTTTTCGAGGTTTTCGAGG 409
QY 361 TCAGAAATTTCTGAAGACAAGTTTGTGCTAAAGTTAAGATGAGATCGTGGTTAGCAAA 420
Db 410 TCAGAAATTTCTGAAGACAAGTTTGTGCTAAAGTTAAGATGAGATCGTGGTTAGCAAA 469
QY 421 GACCAGGTTGAGGATGTTATAGAAAAATCATTTGAGGAGGCAAGAACTCGAGAGATTGGA 480
Db 470 GACCAGGTTGAGGATGTTATAGAAAAATCATTTGAGGAGGCAAGAACTCGAGAGATTGGA 529
QY 481 GACGGCAAGATTTCTTGCTGCCTGTTTCAGATGTAATAAGATCCGCACTGGTGAGCGG 540
Db 530 GACGGCAAGATTTCTTGCTGCCTGTTTCAGATGTAATAAGATCCGCACTGGTGAGCGG 589
QY 541 GGTGATAAGCTGAGAGGATGACAGGAGGCGGATCTGACATGAGTACTTCTGCT 594
Db 590 GGTGATAAGCTGAGAGGATGACAGGAGGCGGATCTGACATGAGTACTTCTGCT 643

RESULT 3
AAF58583
ID AAF58583 standard; cDNA; 588 BP.
XX
AC AAF58583;
XX

DT 23-APR-2001 (first entry)
XX Arabidopsis thaliana P-PII cDNA fragment.
DE Arabidopsis thaliana; PII; plant nitrogen regulatory gene; P-PII;
XX nitrogen assimilation; transgenic plant; herbicide screening; ss.
KW Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX US6177275-B1.
XX 23-JAN-2001.
XX 23-JUL-1997; 97US-00899330.
XX 24-JUL-1996; 96US-0022328P.
XX (UYNV) UNIV NEW YORK STATE.
XX Coruzzi GM, Lam H, Hsieh M;
PI WPI; 2001-158572/16.
DR P-PSDB; AAB69495.
XX Novel P-PII genes capable of regulating plant nitrogen assimilation,
PT useful for transgenic plant production, and as probes for isolating
PT additional genomic clones having P-PII gene promoters.
XX Claim: 2; Col 37-38; 35pp; English.
XX The present sequence encodes a nitrogen regulatory PII protein. Novel
CC plant PII (also called P-PII) nucleotide sequences have been isolated.
CC They are useful for regulating nitrogen assimilation in plants, and in
CC transgenic plant production. They are also used to engineer organisms
CC that overexpress wild-type or mutant P-PII regulatory proteins. P-PII
CC proteins are useful for in vitro screening of herbicides. P-PII
CC nucleotides may be used as probes for isolating additional genomic clones
CC with the promoters of P-PII genes. P-PII promoters are light- and/or
CC sucrose-inducible, and are suitable for genetic engineering of plants
XX Sequence 588 BP; 164 A; 105 C; 147 G; 172 T; 0 U; 0 Other;
SQ Query Match 48.0%; Score 285; DB 4; Length 588;
Best Local Similarity 70.3%; Pred. No. 2.5e-75;
Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1;
QY 7 GAGGCTACTGCGAACTGGGCTTGCTCACTCCTCTTCAATCTAATAACATCAAGAA 66
Db 7 GCGTCAATGACGAACCCCATCTCAATAACTTCTCGGTTCTATTCTGATCGAAAGAAC 66
QY 67 TTCCTGTTTGTGATTCAGTTTGTGTTGTCAGAGCTTAGACATTCCTCGTTTCTCAC 126
Db 67 ATTGCTTCTGATTCGATTCGATTTGTTCTGATTCAGACATTCCTCGGATCTTGC 126
QY 127 TTAAACACCGGTCGAGCGCGTAAGATANGCC---CCCGTCTGTTCTGTTGATTAATGCC 183
Db 127 CTCGATTTGGTCACAAAGTCACCGAGTAATAACAGTCGTGTTTACCTGCTGTTAGTCC 186
QY 184 CAAAGCTCGCCTGACTACATTCCTGATGCTAAATTTACAAAGTGAAGCAATTCAGG 243
Db 187 CAAATATCTTCTGATATATATCCAGACTCGAAATTTTACAAGTGAAGCAATTCAGA 246
QY 244 CCTGGCGAGTCTCGCAAGTTTCTCGGCTTTGCTAATAAATGTTATTCGAGGTGTACT 303
Db 247 CCATGGAGAATCCAGCAAGTTTTCATCGGCTTTACTGAAATCGGATTCGAGGTGTACT 306
QY 304 GTTCTGATGTTTCGAGGTTTGGTGCTCAAGTGTTCAACTGAGAGGCGGCGGCTCA 363
Db 307 GTTCTGATGTGAGAGGTTTGGTGCACAAAGGAGGTCTACCGAGAGACACGGTGTCT 366
QY 364 GAATTTTCTGAAGACAAGTTTGTGCTAAAGTTAAGATGGAGATCGTGGTTAGCAAGAC 423
Db 367 GAGTTTCTCGGAAGACAATAATTTGTTGCTAAAGTTAAGATGGAAATCGTTGTTAAGAAAGAC 426

QY 424 CAGTTGAGGATGTTATAGAAAAAATCATTGAGGAGCCCAAGAACTGGAGATTGGAGAC 483
Db 427 CAAGTGAATCTGTAATCAACACAATAATTGAAGGAGCAAGGACAGGAGATTGGTGT 486
QY 484 GGCAAGATTTTCTGCTGCCTGTTTTCAGATGTAATAAGACTCCGACTGGTGGGGGT 543
Db 487 GGCAAGATTTTCTGCTGCTGTTTTCAGATGTCATAGAGTTAGGACAGGTGAGCGTGG 546
QY 544 GATAAGGCTGAGAGGATGACAGGAG 568
Db 547 GAGAAAGCAGAGAGAGTACTGGTG 571
RESULT 4
ABZ14328
ID ABZ14328 standard; DNA; 591 BP.
XX AC ABZ14328;
XX DT 21-JAN-2003 (first entry)
XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2133.
XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX OS Arabidopsis thaliana.
XX PN WO200216655-A2.
XX PD 28-FEB-2002.
XX PF 24-AUG-2001; 2001WO-US026685.
XX PR 24-AUG-2000; 2000US-0227866P.
XX PR 26-JAN-2001; 2001US-0264647P.
XX PR 22-JUN-2001; 2001US-0300111P.
XX PA (SCRI) SCRIPPS RES INST.
XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX PI Harper JF, Kreps J, Wang X, Zhu T;
XX WPI; 2002-304127/34.
XX PT Identifying a stress condition to which a plant cell has been exposed and
XX producing plants with increased tolerance to these abiotic stresses.
XX PS Claim 144; SEQ ID NO 2133; 577pp + Sequence Listing; English.
XX CC The invention relates to identifying a stress condition to which a plant
XX cell has been exposed, comprising: (a) contacting nucleic acid
XX representative of expressed polynucleotides in the plant cell with an
XX array or probes representative of the plant cell genome; and (b)
XX detecting a profile of expressed polynucleotides in the plant cell
XX characteristic of a stress response. The method is useful in the
XX production of transgenic plants, cells and seeds and in producing plants
XX with increased tolerance to abiotic stress. The present sequence is that
XX of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
XX in methods of the invention. Note: The sequence data for this patent is
XX not represented in the printed specification but is based on sequence
XX information supplied to Derwent by the European Patent Office
SQ Sequence 591 BP; 165 A; 105 C; 148 G; 173 T; 0 U; 0 Other;
Query Match 48.0%; Score 285; DB 6; Length 591;
Best Local Similarity 70.3%; Pred. No. 2.5e-75;
Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

QY	67	TTCCCTGTTTTGATTTCAAGTTTGTGTCCAGAGCTTATAGCAATCTCTCGGTTTTCTCTCAC	126
DB	67	ATTGCTTTCTCTGATTCGATTTGTTCTGGATTCAGACATTCCTCCGACCATCTTGC	126
QY	127	TTTAAACACCGCGGTCAAGCGGTAAAGATATGCC--CCCCTCGTTCTCTGTGATTAATGCC	183
DB	127	CTCGATTTGGTTCACAAAGTCACCGAGTAATAACAGTCGTGTTTTACCTGTCGTTAGTGCC	186
QY	184	CAAAAGTCGCCCTGACTACATTCCTGATGCTAAATTTCTACAAAGTGGAAAGCAATTTCTCAGG	243
DB	187	CAAATATCTTCTGATTATATTCAGACTCGAAATTTTACAAAGGTGGAAAGCAATTTGTCTAGA	246
QY	244	CCCTGGCGAGTCTCGCAAGTTTCCTCGGCTTTTGCTAAAAATTTGGTATTTTCGAGGTGTTACT	303
DB	247	CCATGGAGAAATCCAGCAAGTTTTCATCGGCTTTTACTGAAATTCGGGATTCGAGGTGTTACT	306
QY	304	GTTTCTGATGTTTCGAGGTTTTGGTGCTCAAGGTGGTTCAACTGAGAGCGAGGCGGGCTCA	363
DB	307	GTTTCTGATGTGAGAGGTTTTGGTGACAAAGGAGGTTCTACCGAGAGACACGGTGGCTCT	366
QY	364	GAATTTTCTGAAGACAAGTTTGTTCCTAAAGTTAAGATCGAGATCGTGTGTAGCAAAAGAC	423
DB	367	GAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGGAAATCGTTGTTAAGAAAGAC	426
QY	424	CAGGTTGAGGATGTTATAGAAAAAATCATTTGAGGAGCGCAAGAACTGGAGAGATTGGAGAC	483
DB	427	CAAGTGGATCTGTAATCAACACAAATAATTGAAGGACCAAGGACAGGAGAGATTTGGTGAT	486
QY	484	GGCAAGATTTTCTTGCTGCCCTGTTTCAGATGTATAAGAGTCGGCACTGGTGAAGCGGGGT	543
DB	487	GGCAAGATTTTGTGTTTGGCTGTGTGCAGATGTATAGAGTTAGGACAGGTGACGCGTGGG	546
QY	544	GATAAGGCTGAGAGGATGACAGGAG	568
DB	547	GAGAAAGCAGAGAAGATGACTGGTG	571

RESULT 5

AAF58581
IC AAF58581 standard; cDNA; 817 BP.

23-APR-2001 (first entry)

xx DE Arabidopsis thaliana P-PII cDNA.

Arabidopsis thaliana; PII; plant nitrogen regulatory gene; P-PII; KW
nitrogen assimilation; transgenic plant; herbicide screening; ss.

OS Arabidopsis thaliana.

PN US6177275-B1.

ED 23-JAN-2001.

23-JUL-1997; 97US-00899330.

24--JUL-1996: 96US-0022328P.

PA (UFTY) UNIV NEW YORK STATE.

PI Coruzzi GM, Lam H, Hsieh M;

DR WPI; 2001-158572/16.

DR P-PSDB; AAB69495.

PT Novel P-II genes capable of regulating plant nitrogen assimilation,
PT useful for transgenic plant production, and as probes for isolating
PT additional genomic clones having P-II gene promoters.

PS Claim 2; Fig 12; 35pp; English.

The present sequence encodes a nitrogen regulatory PII protein. Novel plant PII (also called P-II) nucleotide sequences have been isolated. They are useful for regulating nitrogen assimilation in plants, and in transgenic plant production. They are also used to engineer organisms that overexpress wild-type or mutant P-II regulatory proteins. P-II proteins are useful for in vitro screening of herbicides. P-II nucleotides may be used as probes for isolating additional genomic clones with the promoters of P-II genes. P-II promoters are light- and/or sucrose-inducible, and are suitable for genetic engineering of plants

RESULT 6

AAV21209 01

Continuation (2 of 17) of AAV21209 from base 100001 (Methanococcus jannaschii circular
 wp sequence split into 17 fragments LOCUS AAV21209 Accession Aav21209

WP	Fragment Name	Begin	End
WP	AAV21209_00	1	110000
WP	AAV21209_01	100001	210000
WP	AAV21209_02	200001	310000
WP	AAV21209_03	300001	410000
WP	AAV21209_04	400001	510000
WP	AAV21209_05	500001	610000
WP	AAV21209_06	600001	710000
WP	AAV21209_07	700001	810000
WP	AAV21209_08	800001	910000
WP	AAV21209_09	900001	1010000
WP	AAV21209_10	1000001	1110000

Db 13013 AATGGTATCACCGTCACTGAAGTCAAGGGCTTTGGTCGCCAAAAGGTCATACCGAGATG 13072

QY 352 CAGGGGGGCTCAGAAATTTCTGAAGACAAAGTTTGTGCTAAGTTAAGATGGAGATCGTG 411

Db 13073 TATCGTGGGGCGGAATATGTGGTTGAT---TTTTCACCAAAAATTAAATTTGAGATAGCA 13129

QY 412 GTTAGCAAAAGACCAGGTTGAGGATGTTATAGAAAAATCATTTGAGGAGGCAAGAACTGGA 471

Db 13130 TGTGCTGATGAGATGGTTGATTCAATATTGAGTCAATCATTTAAAGTTGCCAATACAGGT 13189

QY 472 GAGATTGGAGACGGCAAGATTTTCTTGCTGCTGCTGCTTTCAGATGTAATAGAGTCCGCACT 531

Db 13190 AAAATTGGTGATGGTAAAGATTTTGTAGTCCGCTTGAGCGGTGTCATTCGCAATTCGAAC 13249

QY 532 GGTGAGCGGGTGATAAGGC 551

Db 13250 GGCGAATTTGATGAAAGTGC 13269

RESULT 9

ADA30024

ID ADA30024 standard; DNA; 375 BP.

XX

AC ADA30024;

XX

DT 20-NOV-2003 (first entry)

XX

DE DNA encoding Acinetobacter baumannii protein #1311.

XX

KW ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;

KW vaccine; plant biocontrol agent.

XX

OS Acinetobacter baumannii.

XX

PN US6562958-B1.

XX

PD 13-MAY-2003.

XX

PF 04-JUN-1999; 99US-00328352.

XX

PR 09-JUN-1998; 98US-0088701P.

XX

PA (GENO-) GENOME THERAPEUTICS CORP.

XX

PI Breton G, Bush D;

XX

DR WPI; 2003-576092/54.

DR P-PSDB; ADA34150.

XX

New Acinetobacter baumannii proteins and nucleic acids, useful as reagents for diagnosing a bacterial disease, as components of antibacterial vaccines, as targets for antibacterial drugs, to detect the presence of A. baumannii and other Acinetobacter species in a sample, in screening compounds for the ability to interfere with the A. baumannii life cycle or to inhibit A. baumannii infection, and as biocontrol agents for plants. The present sequence represents DNA encoding an A. baumannii protein.

Sequence 375 BP; 122 A; 64 C; 93 G; 96 T; 0 U; 0 Other;

XX

PS Example; SEQ ID NO 1311; 328pp; English.

XX

CC The invention relates to isolated Acinetobacter baumannii nucleic acids. The A. baumannii nucleic acids and polypeptides are useful as reagents for diagnosing a bacterial disease, as components of antibacterial vaccines, as targets for antibacterial drugs, to detect the presence of A. baumannii and other Acinetobacter species in a sample, in screening compounds for the ability to interfere with the A. baumannii life cycle or to inhibit A. baumannii infection, and as biocontrol agents for plants. The present sequence represents DNA encoding an A. baumannii protein.

Query Match 14.1%; Score 83.6; DB 8; Length 375;

Best Local Similarity 55.9%; Pred. No. 9.5e-15;

Matches 180; Conservative 0; Mismatches 139; Indels 3; Gaps 1;

QY 232 GCAATTCTCAGGCCCTGGCGAGTCTCGCAAGTTTCTCGGCTTTGCTAAAAATGGTATT 291

Db 52 GCATTTGTAACACCGTTTAAATTTGATGATGTCGTTGAAGCACTCTCTGACATTTGGTGA 111

QY 292 CGAGGTGTTACTGTTTCTGATGTTCCAGGTTTGGTGTCTCAAGGTGTTCAACTGAGAGG 351

Db 112 CAAGGATTAACCGTAACCTAAGTTAAAGTTTGGTGTCTCAAAAGGACATACAGAACTT 171

QY 352 CAGGGCGGCTCAGAAATTTTCTGAAGACAAAGTTTGTGCTTAAAGTTAAGATGGAGATCGTG 411

Db 172 TACCGCGGCGCTGAGTATGTGGTTGAT---TTCTTACCTAAAGTAAAAATCGAAATTCG 228

QY 412 GTTAGCAAAAGACCAGGTTGAGGATGTTATAGAAAAATCATTTGAGGAGGCAAGAACTGGA 471

Db 229 ATTAGTGATGAATGGTCGACGCGGTAATTTTGTGACTAATCTGGAACAAAGTCATCCGTATCCGTACA 288

QY 472 GAGATTGGAGACGGCAAGATTTTCTTGCTGCTGCTGCTTTCAGATGTAATAGAGTCCGCACT 531

Db 289 AAAATCGGCGAGCGTAAAGATTTTGTGACTAATCTGGAACAAAGTCATCCGTATCCGTACA 348

QY 532 GGTGAGCGGGTGATAAGGCTG 553

Db 349 GGTGAAACAGGACCAGATGCTG 370

RESULT 10

ACF72027

ID ACF72027 standard; DNA; 339 BP.

XX

AC ACF72027;

XX

DT 20-NOV-2003 (first entry)

XX

DE Photorhabdus luminescens nucleotide sequence #10494.

XX

KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;

KW detection; food; gene expression; plant; animal; microorganism; toxin;

KW antibiotic; biopesticide; virulence factor; disease model; plague;

KW whooping cough; gene; ds.

XX

OS Photorhabdus luminescens.

XX

PN WO200294867-A2.

XX

PD 28-NOV-2002.

XX

PF 07-FEB-2002; 2002WO-IB003040.

XX

PR 07-FEB-2001; 2001FR-00001659.

XX

PA (INSP) INST PASTEUR.

PA (CNRS) CNRS CENT NAT RECH SCI.

XX

PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;

PI Buchrieser C;

XX

DR WPI; 2003-148459/14.

XX

CC Genomic sequence of Photorhabdus luminescens and encoded polypeptides, useful e.g. as therapeutic antimicrobials and agricultural pesticides.

PS Claim 2; SEQ ID NO 10494; 1205pp; French.

XX

CC The invention relates to the isolation of genes and their encoded proteins from Photorhabdus luminescens. The isolated sequences are sources of probes and primers for detecting the genome of P. luminescens and related species; to study polymorphisms; for gene analysis and for detection/amplification of the genes. Antibodies (Ab) raised against the polypeptides encoded by the genes are used for detection/identification of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that carry a gene-containing vector are used to select compounds that modulate, regulate, induce or inhibit expression of the genes in plants, animals or microorganisms other than P. luminescens and are able to alter response or sensitivity to toxins and antibiotics produced by P.

Query Match 13.7%; Score 81.4; DB 7; Length 110000;
Best Local Similarity 55.0%; Pred. No. 5.5e-13;
Matches 182; Conservative 0; Mismatches 146; Indels 3; Gaps 1;

QY 221 ACAAGTGGAGCAATTCTCAGGCCCTGGCGAGTCTCGCAAGTTTCCTCGGCTTTGCTAA 280
DB 88821 AAAAGATTGATGCGATTATCAAAACCTTTCAAAATAGATGATGCGTGAAGCTCTGGCGG 88880

QY 281 AAATTGGTATTCGAGGTGTTACTGTTTCTGATGTTTCGAGGTTTTGGTGCTCAAGGTGGTT 340
DB 88881 AAGTGGGTATCACCGGATGACAGTAACAGAGGTGAAAGGTTTTTGGCGCCAAAAGGTC 88940

QY 341 CAACTGAGAGCGCGGCTCAGAAATTTTCTGAAGACAAAGTTTGTGCTAAAGTTAAGA 400
DB 88941 ATACAGAGCTGTATCGGGTGCGAATATATGTTGGAT---TTTCTGCCAAAAGTGAAA 88997

QY 401 TGGAGATCGTGGTTAGCAAGACCAGGTTGAGGATGTTATAGAAAAATCATTCAGGAGG 460
DB 88998 TAGAAATTGTGTCGCCAGATGATATTGTGATACCTGTGTTGAAACCATATTCAGACGG 89057

QY 461 CAAGAACTGGAGAGATTGGAGACGGCAAGATTTTCTGCTGCTGCTGTTTCAGATGTAATA 520
DB 89058 CACAGACCGGAAAAATCGGTGTTGTTAAATATTGTTATTTGATGTAGCAGGTGTTGTC 89117

QY 521 GAGTCCGCACTGGTGAGCGGGGTGATAAGGC 551
DB 89118 GTATCCGCACTGGTGAGCAGGATGAAGAGGC 89148

RESULT 13
AAT42063_03
Continuation (4 of 19) of AAT42063 from base 300001 (Haemophilus influenzae complete genome)
WP Sequence split into 19 fragments LOCUS AAT42063 Accession Aat42063

Fragment Name	Begin	End
WP AAT42063_00	1	110000
WP AAT42063_01	100001	210000
WP AAT42063_02	200001	310000
WP AAT42063_03	300001	410000
WP AAT42063_04	400001	510000
WP AAT42063_05	500001	610000
WP AAT42063_06	600001	710000
WP AAT42063_07	700001	810000
WP AAT42063_08	800001	910000
WP AAT42063_09	900001	1010000
WP AAT42063_10	1000001	1110000
WP AAT42063_11	1100001	1210000
WP AAT42063_12	1200001	1310000
WP AAT42063_13	1300001	1410000
WP AAT42063_14	1400001	1510000
WP AAT42063_15	1500001	1610000
WP AAT42063_16	1600001	1710000
WP AAT42063_17	1700001	1810000
WP AAT42063_18	1800001	1830121

Query Match 13.1%; Score 77.6; DB 2; Length 110000;
Best Local Similarity 55.1%; Pred. No. 7.7e-12;
Matches 174; Conservative 0; Mismatches 139; Indels 3; Gaps 1;

QY 221 ACAAGTGGAGCAATTCTCAGGCCCTGGCGAGTCTCGCAAGTTTCCTCGGCTTTGCTAA 280
DB 63919 AAAAATCGAAGCAATGATTAAACCCCTTAAATTAGACGATGTCGAGAAAGTCTTTTCAG 63978

QY 281 AAATTGGTATTCGAGGTGTTACTGTTTCTGATGTTTCGAGGTTTTGGTGCTCAAGGTGGTT 340
DB 63979 ATATTGGTATTCAGGTATGACAATCACAGAAGTACCGGATTTGGTCTGTCAAAAGGTC 64038

QY 341 CAACTGAGAGCGAGCGGCTCAGAAATTTTCTGAAGACAAAGTTTGTGCTAAAGTTAAGA 400
DB 64039 ATACAGAACTTTATCGTGGTGCGGAATATATGTTGGAT---TTTCTGCCAAAAGTGAAAT 64095

QY 401 TGGAGATCGTGGTTAGCAAGACCAGGTTGAGGATGTTATAGAAAAATCATTCAGGAGG 460

DB 64096 TGAAGTGGTAGTTCTCTGATGAGCTTGTGGATCAATGTATTGAAGCGATTATTGAACGG 64155

QY 461 CAAGAACTGGAGAGATTGGAGACGGCAAGATTTCTTGTGCTGCTGTTTTCAGATGTAATA 520
DB 64156 CACAAACAGGTAAATCGGTGACGGCAAAATTTTGTATTATCAGTTGAGAGAGCCATCC 64215

QY 521 GAGTCCGCACTGGTGA 536
DB 64216 GCATTGCGACGGCGA 64231

RESULT 14
ABZ40022
ID ABZ40022 standard; DNA; 363 BP.
XX AC ABZ40022;
XX DT 07-MAR-2003 (first entry)
XX DE N. gonorrhoeae nucleotide sequence SEQ ID 4633.
XX KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX OS Neisseria gonorrhoeae.
XX PN WC200279243-A2.
XX PD 10-OCT-2002.
XX PF 12-FEB-2002; 2002WO-IB002069.
XX PR 12-FEB-2001; 2001GB-00003424.
XX PA (CHIR-) CHIRON SPA.
XX PI Fontana MR, Pizza M, Massignani V, Monaci B;
XX WPI; 2003-058415/05.
XX P-ESDB; ABP79052.
XX PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX PT medicament for treating or preventing N. gonorrhoeae infection.
XX PS Disclosure; Page 528; 815pp; English.
XX CC The present invention relates to proteins from Neisseria gonorrhoeae.
XX CC Also disclosed are the nucleic acid molecules encoding the proteins and
XX CC antibodies that specifically bind to the proteins. The composition
XX CC comprising the protein, nucleic acid or antibody is useful for the
XX CC manufacture of a medicament for treating or preventing N. gonorrhoeae
XX CC infection, this may be in the form of a vaccine or gene therapy.
XX CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
XX CC molecules of the invention
SQ Sequence 363 BP; 84 A; 77 C; 123 G; 79 T; 0 U; 0 Other;

Query Match 13.0%; Score 77; DB 7; Length 363;
Best Local Similarity 54.1%; Pred. No. 9.1e-13;
Matches 180; Conservative 0; Mismatches 150; Indels 3; Gaps 1;

QY 221 ACAAGTGGAGCAATTCTCAGGCCCTGGCGAGTCTCGCAAGTTTCCTCGGCTTTGCTAA 280
DB 32 AAAAATCGAGCGGATGTCAAACCGTTCAAGCTCGACGACGTGCGGAGGCGTTGACGG 91

QY 281 AAATTGGTATTCGAGGTGTTACTGTTTCTGATGTTTCGAGGTTTTGGTGCTCAAGGTGGTT 340
DB 92 AAATCGGCAATTACGGGCATGACCGTCAGCGAGGTCAAAGGTTTCGGCAGGCAAGGGGC 151

QY 341 CAACTGAGAGCGAGCGGCTCAGAAATTTTCTGAAGACAAAGTTTGTGCTAAAGTTAAGA 400
DB 152 ATACGGAAATCTATCGGGTGGGAATACGCCGTCGAT---TTCTGCCCAAGTCAAAA 208

QY 401 TGGAGATCGTGGTTAGCAAGACCAGGTTGAGGATGTTATAGAAAAATCATTCAGGAGG 460

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 19:07:55 ; Search time 1376.09 Seconds
(without alignments)
12890.278 Million cell updates/sec

Title: US-09-756-541-16
Perfect score: 594
Sequence: 1 CGGCACGAGGCTACTGGAA.....CTGACATGAGTACTTCTGCT 594

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	408.8	68.8	542	14	T15252
2	321	54.0	782	14	CB289081
3	300.8	50.6	723	13	BQ798001
4	298.6	50.3	597	10	BE239550

5	298.6	50.3	714	12	BI311072
6	297	50.0	812	14	CA922829
7	293	49.3	612	10	BF645522
8	276.6	46.6	759	14	CP920467
9	275.2	46.3	711	12	BI932336
10	275	46.3	732	9	AJ558383
11	274.2	46.2	648	9	AI773079
12	273.6	46.1	613	9	AW035791
13	273.6	46.1	702	12	BI932123
14	273.6	46.1	736	12	BI932913
15	273.6	46.1	741	10	AW738071
16	273.6	46.1	804	12	BI930336
17	273.6	46.1	810	12	BI930060
18	272.6	45.9	799	12	BM113473
19	272	45.8	630	10	AW929534
20	272	45.8	685	12	BI929601
21	270.4	45.5	756	12	BG595705
22	270.2	45.5	677	10	BE823431
23	269.2	45.3	866	13	C81819
24	267	44.9	719	12	BJ573686
25	266.2	44.8	732	14	CD820903
26	265.4	44.7	657	14	CB349285
27	265.4	44.7	701	14	CB349370
28	264.6	44.5	458	10	BE210153
29	264.6	44.5	669	14	CD822497
30	263	44.3	685	14	CD825321
31	263	44.3	745	14	CD819046
32	262.2	44.1	675	14	CB346213
33	248.4	41.8	549	14	CB261070
34	248.2	41.8	560	14	CA514765
35	246.4	41.5	579	10	AW223643
36	241.2	40.6	673	14	CB085731
37	239.2	40.3	540	14	CB349912
38	238.6	40.2	551	14	CB349990
39	238.6	40.2	598	14	CB350294
40	236.4	39.8	540	10	AW624562
41	233.4	39.3	532	14	CB350222
42	227.6	38.3	564	13	BQ118918
43	227.4	38.3	729	14	CF475242
44	227.2	38.2	796	14	CF395143
45	223.6	37.6	690	14	CA290108

ALIGNMENTS

RESULT 1	T15252	T15252	542 bp	mRNA	linear	EST 23-MAY-2001
LOCUS	crs852	lambdaZAPST	Ricinus communis	cDNA clone	pcrs852	similar to
DEFINITION	nitrogen-regulatory protein,	mRNA sequence.				
ACCESSION	T15252					
VERSION	T15252.1	GI:14190796				
KEYWORDS	EST.					
SOURCE	Ricinus communis (castor bean)					
ORGANISM	Ricinus communis					
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Malpighiales; Euphorbiaceae; Acalyphoideae; Acalyphaeae; Ricinus.					
AUTHORS	1 (bases 1 to 542)					
TITLE	vandeLoe,F.J., Turner,S. and Somerville,C.					
JOURNAL	Expressed sequence tags from developing castor seeds					
COMMENT	Plant Physiol. 108, 1141-1150 (1995)					
	Contact: Somerville CR					
	Carnegie Institution					
	Carnegie Institution, 290 Panama St, Stanford, CA 94305					
	Tel: 4153251521					
	Email: crs@andrew.stanford.edu					
	Seq primer: T3.					
FEATURES	Location/Qualifiers					
source	1. .542					
	/organism="Ricinus communis"					

/mol_type="mRNA"
/strain="Baker 296"
/db_xref="taxon:3988"
/clone="pcrs852"
/clone_lib="lambdaZAP5"
/note="Vector: lambdaZAPII; Site_1: EcoRI; Site_2: XhoI;
Poly(A)+ RNA was purified from developing stage_III to
stage V (Greenwood & Bewley, Can. J. Bot. 60:1751-1760,
1982) endosperm plus embryo of immature castor fruits.
cDNA was synthesized and cloned into lambdaZAPII according
to the instructions of the manufacturer (Stratagene):
synthesis was primed from the poly(A) tail, and cloned
directionally into XhoI (3') and EcoRI (5') sites. In few
cases, sequence data indicated that this directionality
was reversed. Partial cDNA clones predominate."

ORIGIN

Query Match 68.8%; Score 408.8; DB 14; Length 542;
Best Local Similarity 95.5%; Pred. No. 9.4e-103;
Matches 428; Conservative 0; Mismatches 18; Indels 2; Gaps 1;

Qy 10 GCTACTCGAAACTGGCTTGCTCACTCTCTTCAATCTAATAACATCAAGAAAGATTC 69
Dd 1 GCTACTCGAAACTGGCTTGCTCACTCTCTTCAATCTAATAACATCAAGAAAGATTC 60
Qy 70 CCTGTTTGTATTTCAGTTTGTGTTGTCAGAGCTTAGACATTCGCGTTTCTCATT 129
Dd 61 CCTGTTTGTATTTCAGTTTGTGTTGTCAGAGCTTAGACATTCGCGTTTCTCATT 120
Qy 130 AACACCGCGGTCAAGCGCGTAAGATATGCCCGTCCGTTCTGTAATGCCCCAAGC 189
Dd 121 AACACCGCGGTCAAGCGCGTAAGATATGCCCGTCCGTTCTGTAATGCCCCAAGC 180
Qy 190 TCGCCTGACTACATTCCTGATGCTAATTTCTACAAAGTGAAGCAATTCAGGCCCTGG 249
Dd 181 TCGCCTGACTACATTCCTGATGCTAATTTCTACAAAGTGAAGCAATTCAGGCCCTGG 240
Qy 250 CGAGTCTCGCAAGTTTCTCGGCTTTCCTAATAAATTGGTATTCGAGGTGTTCTTCT 309
Dd 241 CGAGTCTCGCAAGTTTCTCGGCTTTCCTAATAAATTGGTATTCGAGGTGTTCTTCT 300
Qy 310 GATGTCGAGGTTTGGTGCTCAAGTGTTCAACTGAGAGCGAGCGGCTCAGAAATT 369
Dd 301 GATGTCGAGGTTTGGTGCTCAAGTGTTCAACTGAGAGCGAGCGGCTCAGAAATT 360
Qy 370 TCTGAAGACAAGTTTGTGCTAAAGTTAAGATGGAGATCGTGTTAGCAAAAGACCAGTT 429
Dd 361 NCTGAAGNCAAGTTTGTGCTAAAGTAAAGATGGAGCCGTTGT--NGCAAAGCCAGGTT 418
Qy 430 GAGGATGTTATAGAAAAAATCAATGAGG 457
Dd 419 GAGGNTGTTATTGGAAATNATTGGGG 446

RESULT 2

CB289081 782 bp mRNA linear EST 27-FEB-2003
LOCUS V-B-112F06 VAN-Baker-1 Vitis aestivalis cDNA clone V-B-112F06 5',
mRNA sequence.

ACCESSION CB289081
VERSION CB289081.1 GI:28602822
KEYWORDS EST.

SOURCE Vitis aestivalis
ORGANISM Vitis aestivalis

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.

REFERENCE 1 (bases 1 to 782)
AUTHORS Hou,H.S., Phanikanth,T.V., Kovacs,L. and Qiu,W.P.
TITLE Expressed sequence tags of young leaf tissues of a
disease-resistant Vitis aestivalis var. Norton
JOURNAL Unpublished (2003)
COMMENT Contact: Wenping Qiu

Department of Fruit Science
Southwest Missouri State University-Mountain Grove
9740 Red Spring Road, Mountain Grove, MO 65711-2999, USA
Tel: 417 926 4105
Fax: 417 926 6646
Email: weq070f@smsu.edu
Insert Length: 782 Std Error: 0.00
Plate: VAN-Baker-1-12 row: F column: 06
Seq primer: T3 PRIMER
High quality sequence stop: 782
POLYA=No.

FEATURES
source

Location/Qualifiers
1..782
/organism="Vitis aestivalis"
/mol_type="mRNA"
/cultiivar="Norton"
/db_xref="taxon:3605"
/clone="V-B-112F06"
/tissue type="Leaf"
/dev_stage="Young leaf"
/lab_host="XL10-Gold E.coli"
/clone_lib="VAN-Baker-1"
/note="Vector: pBluescript II SK (+) Phagemid; Site 1: Xho
I; Site 2: EcoR I; VAN-Baker-1 is a cDNA library of Norton
grape young leaves (Vitis. aestivalis var. Norton). Norton
grapevines were grown under normal greenhouse conditions.
The cDNA synthesis and library construction was performed
according to the instruction manual for pBluescript II XR
cDNA library construction kit provided by Stratagene."

ORIGIN

Query Match 54.0%; Score 321; DB 14; Length 782;
Best Local Similarity 75.4%; Pred. No. 2.8e-78;
Matches 399; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

Qy 62 AAGAAATCCCTGTTTGTGATTTGATTTGATTTGTTTGTCCAGAGCTTAGACATTTCTCGTTT 121
Dd 65 AAGAAATGCTCTCATAGACTGCACTTTGATCCGCCCAATCTCAAGATTTCTGGAATT 124
Qy 122 CTCACCTTTAACACCGGTCAGCGGTCAAGCGGTAAAGATATGCCCGTCTGTTCTGTGATTATG 181
Dd 125 TCCAGTTTAATTTGTGCTAAACCGCTCAAGAAATGGATCTGTTCTTCCCATGTTCAGAG 184
Qy 182 CCCAAAGCTCGCTGACTACATTCCTGATGCTAAATTTCTACAAAGTGGAGCAATTTCTCA 241
Dd 185 CTCAGAGCTCTCCAGATTATACCCAGAGCGCGCAGTTTACAAAGTTGAAGCGATCTCTGA 244
Qy 242 GGCCCTGGGAGTCTCGCAAGTTTCTCGGCTTTGCTAATAAATTTGTAATTCGAGGTGTTA 301
Dd 245 GGCCCTGGGGAATCCAGCAGGTTTCTTCGGCTTTGCTGAAATGGGTATTCGTGTTGTTA 304
Qy 302 CTGTTTCTGATGTTGAGGTTTGTGGTGCTCAAGGTGGTTCAACTGAGAGGCAGGCGGCT 361
Dd 305 CTGTTTCTGATGTTGAGGCTTTGGTGCTCAAGCGGATCGCCTGAAAGACAGGCAGGCT 364
Qy 362 CAGAAATTTCTGAAGACAAGTTTGTGCTTAAAGTTAAGATGGAGATCGTGTAGCAAAG 421
Dd 365 CTGAATTTCTGAGGACAAATTTGTTGCTTAAAGTTAAATGGAGATTGTTGAGCAAAG 424
Qy 422 ACCAGGTTGAGGATGTTATAGAAAAATCATTTGAGGAGCGCAAGAACTGGAGAGATTGGAG 481
Dd 425 ACCAGGTTGAAGCAGTAATTGACAGATCAATGAGGTGGCAAGACTGGAGAGATTGGTG 484
Qy 482 ACGGCAAGATTTTCTGTCGCTGTTTCAGATGTAATAAGAGTCCGCACTGGTGAGCGGG 541
Dd 485 ATGGCAAGATCTTTTGGTGCCCATATCGGATGTGATAAGAGTTCCGACCGTGAACGTG 544
Qy 542 GTGATAAGGCTGAGAGGATGACAGAGGGCGGATCTGACATGAGTACTTC 590
Dd 545 GAGAGATGGCTGAGAGGATGACTGGTGGCGGAAGTATGATATGTCCACTAC 593

RESULT 3


```
Db 28 ATTGGAACCGAAGCTGTTCAACGGTTTGAATTTTCATATCAATGAAGCAATTCCT 87
Qy 73 GTTTTGTGATTTAGTTTGTGTTCCAGAGCTTAGACATTTCTCGGTTTCTCACTTTAAC 132
Db 88 TTTTCAAGCTTCAGGCTCATTCGCAAGCGTTTCGGAGATTC-----TTCTCATCGCAAT 141
Qy 133 ACCGGGTCAAGCGGTAAAGATATGCCCGTCTGTTTCTGATTAATGCCCAAGCTCG 192
Db 142 GTGGTCTCTAAAGCAACGGAAATGATCGATCTTCTCCAAAATCAGAGCTCAAAACCTT 201
Qy 193 CCTGACTACATTCCTGATGCTAAATTTCTACAAGTGAAGCAATTTCTCAGGCCCTGGCGA 252
Db 202 CCTGACTATGTTCTCTGAATCCAAAGTTTACAAAGTTGAAGCCATTTCTCAGGCCATGGAGA 261
Qy 253 GTCTCGCAAGTTTCTCTCGGCTTTGCTAAATAATTTGGTATTTCTGAGGTGTTTCTGAT 312
Db 262 ATCCCTCAGGTTTCTCTCGGTTTGTGAAATGGGAATTCGTGTCTCACTGTATCTGAT 321
Qy 313 GTTCGAGGTTTGTGCTCAAGTGTGTTCAACTGAGAGCGGCGGCTCAGAAATTTTCT 372
Db 322 GTCAAGGGGTTTGTGCTCAGGTTGGCTCAAAAGAGAGCGGAGGCTCCGAATTTTCT 381
Qy 373 GAAGACAAAGTTTGTGCTAAAGTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 432
Db 382 GAAGACAAATTTGTGCTCAAAAGTTAAATGGAATAGTGGTGAAGAAAGACCAAGTTGAG 441
Qy 433 GATGTTATAGAAAATCATTGAGGAGGCAAGAACTGGAGAGATTGGAGAGCGCAAGATT 492
Db 442 GCAGTGATAACAAATATTATGGAGACGGCAAGAACTGGGGAGATTGGTGAAGCAAGATT 501
Qy 493 TTCTTGCTGCTGTTTCTGATGTAATAAGTCCGCACTGGTGAGCGGGTGATGAAGGCT 552
Db 502 TTCTTGATCCCTGATCTGATGTAATAAGAAATCCGACAGGTGAGCGGTGGGAGCAGGCT 561
Qy 553 GAGAGGATGACAGGAGCGGATCTGAC 579
Db 562 GAGAGGATGGCTGGGGGACTAACTGAC 588
```

```
RESULT 5
BI311072 714 bp mRNA linear EST 20-JUL-2001
LOCUS EST5312822 GESD Medicago truncatula cDNA clone pGESD9J4 5' end,
DEFINITION mRNA sequence.
ACCESSION BI311072
VERSION BI311072.1 GI:14985399
KEYWORDS Medicago truncatula (barrel medic)
SOURCE Medicago truncatula
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 714)
AUTHORS Grusak,M.A., Samac,D.A., Town,C.D., Van Aken,S., Utterback,T.,
Cho,J. and Fraser,C.M.
TITLE ESTs from developing reproductive tissues of Medicago truncatula
JOURNAL Unpublished (2001)
COMMENT Contact: Michael A. Grusak
USDA/ARS Children's Nutrition Research Center
Baylor College of Medicine
1100 Bates Street, Houston, TX 77030-2600, USA
Tel: 713-798-7044
Fax: 713-798-7078
Email: mgrusak@bcm.tmc.edu
B397927e
CIGR sequence name: MTPAP50TK
More information is available at: www.medicago.org
Seq primer: SKmod (CTA GAA CTA gtg gat CC).
FEATURES
source location/Qualifiers
1..714
/organism="Medicago truncatula"
```

```
/mol_type="mRNA"
/cultivar="A17"
/db_xref="taxon:3880"
/clone="pGESD9J4"
/tissue_type="immature seeds"
/dev_stage="Immature seeds, 11 to 19 days after
pollination"
/clone_lib="GESD"
/notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Immature seeds, collected from pods ranging in age
from 11 to 19 days after pollination, were harvested from
greenhouse-grown plants. Seeds were removed and
separated from pod walls and were immediately frozen in
liquid nitrogen. Seeds throughout the age range were
pooled for mRNA extraction. cDNA was prepared from polyA+
enriched RNA. The cDNA was directionally ligated into
the Unizap XR vector from Stratagene and packaged using
Gigapack III Gold packaging extracts. Plasmids containing
cDNA inserts were excised from the recombinant lambda-Zap
phage using Ex-assist helper phage and propagated in
XL0LR cells."
```

ORIGIN

```
Query Match 50.3%; Score 298.6; DB 12; Length 714;
Best Local Similarity 71.8%; Pred. No. 4.7e-72;
Matches 407; Conservative 0; Mismatches 154; Indels 6; Gaps 1;

Qy 13 ACTGCGAACTGGGCTGCTCACTCCTCTTCAATTAATCAATCAAGAAAGAAATTCCT 72
Db 27 ATTGCGAAACCGAAGCTGTTCAACGGTTTGAATTTTCATATCAATGAAGCAATTCCT 86
Qy 73 GTTTTGTGATTTCAAGTTTGTGTTTGTCCAGAGCTTAGACATTTCTCGGTTTCTCACTTAAAC 132
Db 87 TTTTCAAGCTTCAGGCTCATTCGCAAGCGTTTCGGAGATTC-----TTCTCATCGCAAT 140
Qy 133 ACCGGGTCAAGCGGTAAAGATATGCCCGCTGTTTCTGATTAATGCCCAAGCTCG 192
Db 141 GTGGTCTCTAAAGCAACGGAAATGCATCGATTTCTCCAAAATCAGAGCTCAAAACCTT 200
Qy 193 CCTGACTACATTCCTGATGCTAAATTTCTACAAGTGAAGCAATTTCTCAGGCCCTGGCGA 252
Db 201 CCTGACTATGTTCTCTGAATCCAAGTTTTCACAAAGTTGAAGCCATTTCTCAGGCCATGGAGA 260
Qy 253 GTCTCGCAAGTTTCTCTCGGCTTGTGCTAAATAATTTGGTATTCGAGGTGTTACTGTTTCTGAT 312
Db 261 ATCCCTCAGGTTTCTCTCGGTTTGTGTAATAATGGGAATTCGTGTGTTCTACTGTATCTGAT 320
Qy 313 GTTCGAGGTTTGTGCTCAAGGTGTTTCAACTGAGAGCGGCGGCTCAGAAATTTTCT 372
Db 321 GTCAAGGGGTTTGTGCTCAGGTTGGCTCAAAAGAGAGCGGAGGCTCCGAATTTTCT 380
Qy 373 GAAGACAAGTTTGTGCTAAAGTTAAGATCGAGATCGTGTAGCAAGACCAAGGTTGAG 432
Db 381 GAAGACAATTTGTGCTCAAGTTAAATGGAATAGTGGTGAGAAAGACCAAGGTTGAG 440
Qy 433 GATGTTATAGAAAATCATTGAGGAGGCAAGCAACTGGAGAGATTGGAGACGGCAAGATT 492
Db 441 GCAGTGATAACAAATATTATGGAGACGGCAAGAACTGGGGAGATTGGTGAATGGCAAAATT 500
Qy 493 TTCTTGCTGCTGTTTCTGATGTAATAAGAGTCCGCACTGGTGAGCGGGGTGATAGGCT 552
Db 501 TTCTTGATCCCTGATCTGATGTAATAAGAAATCCGACAGGTGAGCGGTGGGAGCAGGCT 560
Qy 553 GAGAGGATGACAGGAGGCGGATCTGAC 579
Db 561 GAGAGGATGGCTGGGGGACTAACTGAC 587
```

```
RESULT 6
CA922829/c 812 bp mRNA linear EST 09-MAY-2003
LOCUS CA922829 MTUS Medicago truncatula cDNA clone MTUS-58H7, mRNA
DEFINITION EST640547 MTUS Medicago truncatula cDNA clone MTUS-58H7, mRNA
sequence.
```


5 prime sequence.
 Location/Qualifiers
 1. .648
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultiivar="R11-12 (35S::Pto in Rio Grande x Money Maker)"
 /db_xref=taxon:4081"
 /clone="CLER5E12"
 /tissue_type="leaf"
 /dev_stage="4-week old"
 /lab_host="SOLR"
 /clone_lib="tomato resistant, Cornell"
 /note="Vector: pBluescript SK(-); Site 1: EcoR1; Site 2:
 Xho1; cLER - Tomato Pseudomonas Resistant EST Library.
 Directionally cloned cDNAs inserted into pBluescript
 SK(-) at 5' end with EcoR1 and 3' end with Xho1 site."

ORIGIN

Query Match	46.2%	Score 274.2;	DB 9;	Length 648;
Best Local Similarity	75.7%	Pred. No. 2.9e-65;		
Matches 339;	Conservative	0;	Mismatches 109;	Indels 0; Gaps 0;

```

QY 135 CGCGGTCAAGCGGTAAGATATGCCCCCGTCGTTCTCTGTGATTAATGCCCAAGCTCGCC 194
    ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
Db 181 CACCTTCAAACGCTGTCAAATGCTCCTTCTTCCCGANTATCAGAGCCCAAACTCTCC 240

QY 195 TGACTACATTCCTGATGCTAAATTTCTACAAAGTGGAAGCAATTTCTCAGGCCCTGGCGAGT 254
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 AGATTTCGTCCCTGATGCGAAGTTTTCACAAAGTTGAAGCAATTTCTAAGACCTTGGAGAAT 300

QY 255 CTCGCAAGTTTCCTCGGCTTTGCTAAAAATTSATTCGAGGTGTTACTGTTTCTGATGT 314
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 TCAACAGGTTTCTTCGSCACTACTGAAATGGCATTGCTGGTGTCACTGTTTCGGATGT 360

QY 315 TCGAGGTTTTTGGTGTCTCAAGGTGGTTCAACTGAGAGCGCAGGCGGCTCAGAAATTTCTGA 374
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 TCGTGGTTTTTGGCGCCCAAGGTGGCTTGACTGAGAGCAAGCTGGCTCTGAATTTCTTGA 420

QY 375 AGACAAGTTTGTGTTAAAGTTTAAGATGGAGATCGTGGTTAGCAAAGACCAGGTTGAGGA 434
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 AGACACGTTTGTGTGCAAAAGTTTAAATGGAATTTGTTGTCTCAGCAAAGACCAGGTTGAAG 480

QY 435 TGTATAGAAAAATTCATTGAGGAGGCAAGAACTCGAGAGATTTGGAGACGGCAAGATTTT 494
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 AGTCATTGCCAAGATAATTGAAGAGCGCAAGAACTGTTGAATAGGTGATGGAAGAATATT 540

QY 495 CTTGCTGCCTGTTTCAAGATGTAATAAGAGTCCGCACCTGCTGAGCGGGGTGATAGGCTGA 554
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 CTTGACTCCCATCTCCGATGTTATAAGAGTTCCGCACTGGTGAACGGNGAGAAAGGCAGA 600

QY 555 GAGGATGACAGGAGGCGCATCTGACATG 582
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 601 GAGGATGATGGAGGGCATGCTGATATG 628
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  
```

RESULT 12	
AW035791	
LOCUS	613 bp mRNA linear EST 18-MAY-2001
DEFINITION	EST281945 tomato callus, TAMU Lycopersicon esculentum cDNA clone
	cLEC36D8, mRNA sequence.
ACCESSION	AW035791
VERSION	AW035791.1 GI:5894547
KEYWORDS	EST.
SOURCE	Lycopersicon esculentum (tomato)
ORGANISM	Lycopersicon esculentum
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
	asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
	1 (bases 1 to 613)
REFERENCE	
AUTHORS	Alcala, J., Vrebalov, J., White, R., Matern, A.L., Vision, T., Holt, I.E., Liang, F., Upton, J., Craven, M.B., Bowman, C.L., Ahn, S., Ronning, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.

ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 702)

REFERENCE
AUTHORS van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
Generation of ESTs from tomato flower tissue, buds 8 mm -
preanthesis

TITLE Lycopersicon esculentum (tomato)

JOURNAL
COMMENT Unpublished (2001)
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
This clone is available through the Clemson University Genomics
Institute

FEATURES
source
Seq primer: T3.
Location/Qualifiers
1..702
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOC22G23"
/tissue_type="flower"
/dev_stage="buds 8mm to preanthesis"
/clone_lib="tomato flower, 8 mm to preanthesis buds"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research; Flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

ORIGIN
Query Match 46.1%; Score 273.6; DB 12; Length 702;
Best Local Similarity 75.7%; Pred. No. 4.3e-65;
Matches 339; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 135 CGCGGTCAAGCGGTAAAGATATGCCCGTCTGTTCTGTGATTATGCCCAAAGCTCGCC 194
Db 175 CACCTTCAAACGCTGTCAAATGCTCCTTCTTCCGATTATCAGAGCCCAAACCTCTCC 234

QY 195 TGACTACATTCTCGATGCTAAATTTACAAAGTGAAGCAATTTCTAGGCCCTGGCGAGT 254
Db 235 AGATTTCGTCCTGATCGGAAGTTTACAAAGTTGAAGCAATTTCTAGACCTTGGAGAT 294

QY 255 CTCGCAAGTTTCTCGGCTTTGCTAAAAATTTGGTATTCGAGGTGTTACTGTTCTGTAGT 314
Db 295 TCAACAGGTTTCTTCGGCACTACTGAAATGGGCATTCGTGTGTCACTGTTTCGGATGT 354

QY 315 TCGAGGTTTGGTGCTCAAGGTGTTCAACTGAGAGCGCGCGGCTCAGAAATTTCTGA 374
Db 355 TCGTGGTTTGGCGCCCAAGTGGCTTGACTGAGAGCGCAAGCTGGCTCTGAAATTTCTGA 414

QY 375 AGACAAGTTTGTGTTAAAGTTAAAGATGGAGATCGTGGTTAGCAAGACCAGGTTGAGGA 434
Db 415 AGACACGTTTGTGTGCAAAAGTTAAATGGAATTTGTTGTCAGCAAGACCAGGTTGAAGG 474

QY 435 TGTTATAGAAAATCAITTAGAGCGCAAGAACTGGAGAGATTGGAGACGGCAAGATTTT 494
Db 475 AGTCATTGCCATGATAATTGAAGAGCGCAAGAACTGGTGAATAGGTGATGGAAAGATAAT 534

QY 495 CTTCCTGCCTGTTTCAGATGTATAGAGTCCGCACCTGGTGACGGGGTGATAAGGCTGA 554
Db 535 CTTCGACTCCCATCTCCGATGTTTATAGAGTTCCGACTGGTGACGGGAGAAAGGCAGA 594

QY 555 GAGGATGACAGGAGGGCGGATCTGACATG 582
Db 595 GAGGATGATGGGAGGGCATGCTGATATG 622

RESULT 14
BI932913
LOCUS

DEFINITION EST552802 tomato flower, 8 mm to preanthesis buds Lycopersicon
esculentum cDNA clone cTOC24H17 5' end, mRNA sequence.

ACCESSION BI932913

VERSION BI932913.1 GI:16247385

KEYWORDS EST.

SOURCE Lycopersicon esculentum (tomato)

ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 736)

REFERENCE
AUTHORS van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
Generation of ESTs from tomato flower tissue, buds 8 mm -
preanthesis

TITLE Lycopersicon esculentum (tomato)

JOURNAL
COMMENT Unpublished (2001)
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
This clone is available through the Clemson University Genomics
Institute

FEATURES
source
Seq primer: T3.
Location/Qualifiers
1..736
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOC24H17"
/tissue_type="flower"
/dev_stage="buds 8mm to preanthesis"
/clone_lib="tomato flower, 8 mm to preanthesis buds"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research; Flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

ORIGIN
Query Match 46.1%; Score 273.6; DB 12; Length 736;
Best Local Similarity 75.7%; Pred. No. 4.3e-65;
Matches 339; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 135 CGCGGTCAAGCGGTAAAGATATGCCCGTCTGTTCTGTGATTATGCCCAAAGCTCGCC 194
Db 137 CACCTTCAAACGCTGTCAAATGCTCCTTCTTCCGATTATCAGAGCCCAAACCTCTCC 196

QY 195 TGACTACATTCTCGATGCTAAATTTACAAAGTGAAGCAATTTCTAGGCCCTGGCGAGT 254
Db 197 AGATTTCGTCCTGATGCGGAAGTTTACAAAGTTGAAGCAATTTCTAGACCTTGGAGAT 256

QY 255 CTCGCAAGTTTCTCGGCTTTGCTAAAAATTTGGTATTCGAGGTGTTACTGTTCTGATGT 314
Db 257 TCAACAGGTTTCTTCGGCACTACTGAAATGGGCATTCGTGTGTCACTGTTTCGGATGT 316

QY 315 TCGAGGTTTGGTGCTCAAGGTGTTCAACTGAGAGCGCGCGGCTCAGAAATTTCTGA 374
Db 317 TCGTGGTTTGGCGCCCAAGGTGGCTTGACTGAGAGCGCAAGCTGGCTCTGAAATTTCTGA 376

QY 375 AGACAAGTTTGTGTTAAAGTTAAAGATGGAGATCGTGGTTAGCAAGACCAGGTTGAGGA 434
Db 377 AGACACGTTTGTGTGCAAAAGTTAAATGGAATTTGTTGTCAGCAAGACCAGGTTGAAGG 436

QY 435 TGTTATAGAAAATCAITTAGAGCGCAAGAACTGGAGAGATTGGAGACGGCAAGATTTT 494

Db 437 AGTCATTGCCATGATAATTGAAGAGCGCAAGAACTCGTTGAAATAGGTGATGGAAGATATT 496

Qy 495 CTTGCTGCCTGTTTCAGATGTAATAAGAGTCCGCACTGGTGAGCGGGGTGATAAGGCTGA 554

Db 497 CTTGACTCCCATCTCCGATGTTATAAGAGTTCGCACTGGTGAACGGGAGAAAAGGCAGA 556

Qy 555 GAGGATGACAGGAGGGCGCATCTGACATG 582

Db 557 GAGGATGATGGGAGGGCATGCTGATATG 584

RESULT 15

AW738071

LOCUS

DEFINITION

EST339498 tomato flower buds, anthesis, Cornell University

Lycopersicon esculentum cDNA clone cTOD5K11 5', mRNA sequence.

AW738071

ACCESSION

AW738071.1 GI:7647016

VERSION

KEYWORDS

SOURCE

ORGANISM

Lycopersicon esculentum (tomato)

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 741)

REFERENCE

AUTHORS

van der Hoeven,R.S., Bezzeredes,J.L., Materu,A.L., Holt,I.E.,

Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,

Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and

Tanksley,S.D.

Generation of ESTs from tomato flower tissue, anthesis

Unpublished (1999)

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

FEATURES

source

1..741

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/cultivar="TA496"

/db_xref="taxon:4081"

/clone="cTOD5K11"

/tissue_type="flower"

/dev_stage="anthesis"

/clone_lib="tomato flower buds, anthesis, Cornell

University"

/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:

XhoI; supplier: Tanksley; Flower buds and flowers were

taken from greenhouse plants (4-8 wks old, TA496). They

were immediately frozen in liquid nitrogen and then

size-separated while remaining frozen."

ORIGIN

Query Match 46.1%; Score 273.6; DB 10; Length 741;

Best Local Similarity 75.7%; Pred. No. 4.3e-65;

Matches 339; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy 135 CGCGGTCAAGCGGTAAGATATGCCCCCGTCGTTCTGTGATTAATGCCCAAGCTCGCC 194

Db 144 CACCTTCAACGCTGTCAAAATGCTCCTTCTTCCGATTATCAGAGCCCAAACTCTCC 203

Qy 195 TGACTACATTCCTGATGCTAAATTTCTACAAAGTGAAGCAATTTCTCAGGCCCTGGCGAGT 254

Db 204 AGATTTGTCCTCGTGAAGTTTACAAAGTTGAAGCAATTTCTAAGACCTTGAGAAAT 263

Qy 255 CTGCAAGTTTCTCGGCTTTGCTAAAAAATGGTATTCGAGGTGTTACTGTTCTGATGT 314

Db 264 TCAACAGGTTTCTTCGGCACTACTGAAATGGGCATTCGGGTGTCACGTGTTCCGATGT 323

Qy 315 TCGAGGTTTTGGTGTCAAGGTGTTCAACTGAGAGCGGCGGCTCAGAAATTTCTGA 374

Db 324 TCGTGGTTTTGGGGCCCCAAGGTGGCTTGACTGAGAGGCAAGCTGGCTCTGPAATTTCTGA 383

Qy 375 AGACAAGTTTGTGTTGCTAAAGTTAAGATGGAGATCGTGGTTAGCAAAAGACCAGGTTGAGGA 434

Db 384 AGACACGTTTGTGTTGCAAAAGTTAAAATGGAAATTTGTTGTCAAGCAAGACCAGGTTGAAGG 443

Qy 435 TGTATAGAAAAAATCAATTGAGAGGCAAGAACTGGAGAGATTGGAGACGGCAAGATTTT 494

Db 444 AGTCATTTGCCATGATAATTGAGAGAGGCAAGAACTGGTGAATAGGTGATGGAAGATATT 503

Qy 495 CTTGCTGCCTGTTTTCAGATGTAATAAGAGTCCGCACTGGTGAGCGGGGTGATAAGGCTGA 554

Db 504 CTTGACTCCCATCTCCGATGTTATAAAGAGTTCCGACTGGTGAACCGGGAGAAAAGGCAGA 563

555 GAGGATGACAGGAGGGCGCATCTGACATG 582

564 GAGGATGATGGGAGGGCATGCTGATATG 591

Search completed: May 28, 2004, 00:14:55

Job time : 1377.09 secs